

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Point of Contact:
Jan Delaval
Librarian-Physical Sciences
CM1 1E04 Tel: 308-4498

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) <u>6</u>	STN _____
Searcher Phone #: <u>4498</u>	AA Sequence (#) <u>6</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>11/7</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>11/9</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>✓</u>
Clerical Prep Time: <u>20</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>10</u>	Other _____	Other (specify) _____

Searches for User *mjamroz* (Count = 440)

Queries 391 through 440.

Latest

Prev

Next

Oldest

Edit

Help

Return

Main Menu

Logout

09/06/08, 02:11

S #	Updt	Database	Query	Time	Comment
<u>S440</u>	<u>U</u>	USPT,PGPB	(endoglin same antibod\$) same monoclonal	2001-11-09 15:39:06	
<u>S439</u>	<u>U</u>	USPT,PGPB	endoglin same antibod\$	2001-11-09 15:38:39	
<u>S438</u>	<u>U</u>	USPT,PGPB	((TGF) adj beta adj binding adj protein) same antibod\$	2001-11-09 15:33:19	
<u>S437</u>	<u>U</u>	USPT,PGPB	(TGF) adj beta adj binding adj protein	2001-11-09 15:32:52	
<u>S436</u>	<u>U</u>	USPT	5589375.pn.	2001-11-09 15:31:35	
<u>S435</u>	<u>U</u>	USPT	((protein adj tyrosine adj phosphatase) same antibod\$) and @py<1997) and claim	2001-11-09 12:58:12	
<u>S434</u>	<u>U</u>	USPT	((protein adj tyrosine adj phosphatase) same antibod\$) and @py<1999	2001-11-09 12:56:42	
<u>S433</u>	<u>U</u>	USPT	((protein adj tyrosine adj phosphatase) same antibod\$) and @py<1997	2001-11-09 12:24:41	
<u>S432</u>	<u>U</u>	USPT	(protein adj tyrosine adj phosphatase) same antibod\$	2001-11-09 12:24:12	
<u>S431</u>	<u>U</u>	USPT	5580742.pn.	2001-11-09 12:07:49	
<u>S430</u>	<u>U</u>	USPT	(antibody same phosphotyrosine same phosphatase) and @py<1997	2001-11-09 11:59:41	
<u>S429</u>	<u>U</u>	USPT	antibody same phosphotyrosine same phosphatase	2001-11-09 11:59:24	
<u>S428</u>	<u>U</u>	USPT	(antibody same phosphotyrosine same (kinase or phosphatase)) and @py<1997	2001-11-09 11:58:53	
<u>S427</u>	<u>U</u>	USPT	(antibody same phosphotyrosine same (kinase or phosphatase)) same @py<1997	2001-11-09 11:58:42	
<u>S426</u>	<u>U</u>	USPT	antibody same phosphotyrosine same (kinase or phosphatase)	2001-11-09 11:57:52	
<u>S425</u>	<u>U</u>	USPT	5863781.pn.	2001-11-09 10:22:54	



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 19:25:30 ; Search time 438.73 Seconds

(without alignments)
3293.145 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301
Sequence: 1 agagccctgctactggaag.....caatgaatcatgaccgaag 2301

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Archived: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

N_Geneseq_0601.*
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	2301	100.0	2301	AAA29055	Human TGF-beta bin
2	2299.4	99.9	2301	AAA29061	Mutant human TGF-b
3	2299.4	99.9	2301	AAA29062	Human TGF-beta bin
4	2287.8	99.9	2301	AAA29056	Human TGF-beta bin
5	2280.6	99.1	2329	AAA1073	Human secreted pro
6	2047.8	83.0	21	AAA94049	Human DAN/Cerberus
7	2047.8	89.0	9301	AAA29064	Human TGF-beta bin
8	642	27.9	642	AAA94051	Human DAN/Cerberus
9	613.2	26.6	642	AAA29057	Human DAN/Cerberus
10	536.2	23.3	674	AAA29059	Human DAN/Cerberus
11	502	21.8	638	AAA29058	Human DAN/Cerberus

12	467.8	20.3	532	21	AAA29060	Bovine TGF-beta b1
13	422	18.3	1104	21	AAA94050	Human DAN/Cerberus
14	372.2	16.2	35828	21	AAA29063	Human TGF-beta b1
15	193.8	8.4	198	20	AAA40842	Human secreted pro
16	193.4	8.4	196	21	AA05741	Human secreted pro
17	100.4	4.4	936	22	AAE58252	Oligonucleotide D1
18	100.4	4.4	936	22	AAE58254	Oligonucleotide D1
19	100.4	4.4	936	22	AAE58257	Oligonucleotide D1
20	100.4	4.4	936	22	AAE58259	Oligonucleotide D2
21	100.4	4.4	936	22	AAE58262	Oligonucleotide D2
22	100.4	4.4	936	22	AAE58265	Oligonucleotide D1
23	98.6	4.3	936	22	AAE58255	Oligonucleotide D1
24	98.6	4.3	936	22	AAE58257	Oligonucleotide D1
25	98.6	4.3	936	22	AAE58257	Oligonucleotide D1
26	98.6	4.3	936	22	AAE58259	Oligonucleotide D2
27	98.6	4.3	936	22	AAE58255	Oligonucleotide D1
28	98.6	4.3	936	22	AAE58255	Oligonucleotide D1
29	82.6	3.6	954	22	AAE58255	Oligonucleotide D1
30	82.6	3.6	962	21	AAE58255	Oligonucleotide D1
31	82.6	3.6	962	21	AAE58255	Oligonucleotide D1
32	82.6	3.6	962	22	AAE58255	Oligonucleotide D1
33	82.6	3.6	962	22	AAE58255	Oligonucleotide D1
34	69.6	3.0	900	18	AAE58255	Oligonucleotide D1
35	69.6	3.0	900	19	AAE58255	Oligonucleotide D1
36	69.6	3.0	1692	21	AAE58255	Oligonucleotide D1
37	69.6	3.0	1710	20	AAE58255	Oligonucleotide D1
38	69.6	3.0	1719	21	AAE58255	Oligonucleotide D1
39	69.6	3.0	1774	20	AAE58255	Oligonucleotide D1
40	66.4	2.5	114955	20	AAE58255	Oligonucleotide D1
41	57.8	2.5	1412	22	AAE58255	Oligonucleotide D1
42	57.4	2.5	244	22	AAE58255	Oligonucleotide D1
43	55.6	2.4	114955	20	AAE58255	Oligonucleotide D1
44	53	2.3	244	22	AAE58255	Oligonucleotide D1
45	52.4	2.3	2943	17	AAE58255	Oligonucleotide D1

ALIGNMENTS

RESULT 1	
AAA29055	
ID AAA29055 standard; cDNA; 2301 BP.	
NC AAA29055;	
DT 12-SEP-2000 (first entry)	
XX Human TGF-beta binding protein (BEER) cDNA.	
DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KM BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;	
KW bone mineralization; ss.	
XX Homo sapiens.	
OS Key	Location/Qualifiers
FT CDS	48..689
FT	/*tag= a
FT	/product= TGF-beta_binding_protein
PN MO200032773-A1.	
XX 08-JUN-2000.	
PD 24-NOV-1999;	99WO-US27990.
XX 27-NOV-1998;	98US-0110283.
PR (DARW-) DARWIN DISCOVERY LTD.	
PA Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeppe BW;	
PI Van Ness J, Winkler DG;	
XX	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:34:26 : Search time 222.63 seconds
(without alignments)
2340.771 Million cell updates/sec

Title: US-09-668-021-5
Perfect score: 2301
Sequence: 1 agagcctgtctactctgaag.....caatgaatcatgaccgaag 2301

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

arched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82.6	3.6	962	3	US-09-188-930-40 Sequence 40, Appl
2	82.6	3.6	962	3	US-09-188-930-209 Sequence 209, App
3	69.6	3.0	900	1	US-08-468-847B-1 Sequence 1, Appl
4	52.4	2.3	2943	1	US-08-062-747A-7 Sequence 7, Appl
5	49.6	2.2	4257	2	US-08-690-473-1 Sequence 1, Appl
6	49.6	2.2	4257	4	US-09-259-821A-1 Sequence 1, Appl
7	49.6	2.2	4257	4	US-08-843-659-1 Sequence 1, Appl
8	49.6	2.2	12001	1	US-08-458-568A-11 Sequence 11, Appl
9	48.8	2.1	71989	4	US-09-443-501A-2 Sequence 2, Appl
10	48.8	2.1	7218	1	US-08-232-463-14 Sequence 14, Appl
11	47.6	2.0	2150	2	US-08-318-837-1 Sequence 1, Appl
12	46.8	2.0	1910	4	US-09-593-711A-3 Sequence 3, Appl
13	46.8	2.0	1914	1	US-07-601-094-1 Sequence 1, Appl
14	46.8	2.0	1914	1	US-08-012-735-1 Sequence 1, Appl
15	46.6	2.0	4403765	4	US-09-103-840A-2 Sequence 2, Appl
16	44.2	1.9	33529	4	US-09-144-085-3 Sequence 3, Appl
17	43.4	1.9	8438	1	US-07-945-283-1 Sequence 1, Appl
18	43.4	1.9	34303	2	US-08-735-609-4 Sequence 4, Appl
19	43.4	1.9	34303	2	US-08-735-609-4 Sequence 4, Appl
20	43.4	1.9	34303	3	US-09-315-972-4 Sequence 4, Appl
21	43.4	1.9	34303	3	US-09-244-752-4 Sequence 4, Appl
22	43.4	1.9	34303	3	US-09-245-497-4 Sequence 4, Appl
23	43.4	1.9	34382	2	US-08-314-483-6 Sequence 6, Appl
24	43.4	1.9	35408	4	US-08-973-334-3 Sequence 3, Appl
25	43.4	1.9	35408	4	US-09-563-869A-3 Sequence 3, Appl
26	43.4	1.9	35408	4	US-08-549-489-3 Sequence 3, Appl
27	43.4	1.9	35935	2	US-08-735-609-1 Sequence 1, Appl

C	28	43.4	1.9	35935	2	US-08-735-609-1	Sequence 1, Appl
C	29	43.4	1.9	35935	3	US-08-379-452-43	Sequence 43, Appl
C	30	43.4	1.9	35935	3	US-09-315-372-1	Sequence 1, Appl
C	31	43.4	1.9	35935	3	US-09-244-752-1	Sequence 1, Appl
C	32	43.4	1.9	35935	3	US-09-245-497-1	Sequence 1, Appl
C	33	43.4	1.9	35935	3	US-09-409-670-43	Sequence 43, Appl
C	34	43.2	1.9	1578	1	US-08-681-129-1	Sequence 1, Appl
C	35	42.8	1.9	2634	1	US-08-196-218-31	Sequence 31, Appl
C	36	42.8	1.9	13842	4	US-08-681-953-31	Sequence 31, Appl
C	37	42.8	1.9	36778	4	US-09-105-537-30	Sequence 30, Appl
C	38	42.8	1.9	38506	3	US-09-320-878-19	Sequence 19, Appl
C	39	42.8	1.9	835	1	US-08-417-822A-1	Sequence 1, Appl
C	40	42.6	1.9	1269	1	US-08-396-218-1	Sequence 1, Appl
C	41	42.6	1.9	1269	1	US-08-760-116-1	Sequence 1, Appl
C	42	42.6	1.9	20235	1	US-07-642-734C-3	Sequence 3, Appl
C	43	42.6	1.9	20235	3	US-08-439-009A-3	Sequence 3, Appl
C	44	42.6	1.9	5173	1	US-08-242-677-1	Sequence 1, Appl
C	45	42.2	1.8				

ALIGNMENTS

```
RESULT 1
US-09-188-930-40
; Sequence 40, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Iorna
; APPLICANT: Sleeman, Mathew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-40
```

Query Match	3.6%	Score 82.6;	DB 3;	Length 962;
Best Local Similarity	52.5%	Pred. NO. 1.3e-11;		
Matches 287; Conservative	0;	Mismatches 239;	Indels 21;	Gaps 4;
123	tgccagcgcttcaagatgacgagcaatccgcgagctcgagagctaccccgag	182		
107	tgcttggtctttaaataatgacgacgaatcccttattcatatgtgttaacctgc	166		
183	ccttcacgagcgctcgaggaacaaagacatgacccgagcgagagagcgagct	242		
167	ccggacac---accacagacagacacacccctgtaataagcagagtgagcagat	223		
243	ccccccaccc---cttggagaccacaaagacgtgctcgagtcagctcgccgagctgac	299		
224	ttcagtagcactgtagctgacatgcgaacagtcagtgagtgagtcgagagacgcg	283		
300	ttcaccccttcgtagcagcagatgagcgctgacgagcgacacgcgctacccgagctg	359		
284	ttcaccaaatcattcgtgacgagcagtcagcagcagtcagccctcgaagagctggtg	343		
360	tgcttcgagcagtgagcgccgagcagcagcagcagcagcagcagcagcagcagcagc	413		
344	tgccgagcgagtgcttgcttcccttcgcttcgcttcgagagagagctacgag	403		
414	-----aagtgtagcagcagcagtgagcgccagcttcgctgcagatcccgacgcctacgc	467		

```

Db 404 acaaaactgagagccggagagctctcagagatggcggltgtlctcaacgacaagacgcgc 463
Qy 468 ggcgagcgctgacgctgtctgtcccggtggtgagggccgcgcgcgcgcgcgcgcgcgcgc 527
Db 464 acccagagagatccagctgcgcaggtgtcaggaagcgcgcgcgcgcgcgcgcgcgcgcgcgc 520
Qy 528 ctgtgtgcctctgtcagatgcaagcgccctcaccgcgtctcccaacacagctcgagctcaag 587
Db 521 gtgtlctcagcgctgtcagagtgtaacacccgtctcagcaacagctcgcgcgcgcgcgcgc 580
Qy 588 gacttcggagccgagcgagcgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 647
Db 581 tttaagaagcgtgtcgcagcagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 640
Qy 648 agcgcca 654
Db 641 aaatcca 647

```

RESULT 2 US-09-188-930-209 Sequence 209, Application US/09188930A

```

Patent No. 6150502
GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Onrust, Rene
: APPLICANT: Murison, James Greg
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: TITLE OF INVENTION: and Methods For Their Use
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: NUMBER OF SEQ. ID NOS: 348
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 209
: LENGTH: 962
: TYPE: DNA
: ORGANISM: Mouse
US-09-188-930-209

```

Query Match 3.6%; Score 82.6; DB 3; Length 962;

Best Local Similarity 52.5%; Pred. No. 1.3e-11; Mismatches 239; Indels 21; Gaps 4;

```

Db 123 tggcagcgcttcaagaatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 182
Db 107 tgttgcctttaaataatgtgacacagaatcttattcacaatgtgtttaaactgtc 166
Qy 183 cctccacgcgagctgagaaacaaacacatgaacccgcgcgcgcgcgcgcgcgcgcgcgcgc 242
Db 167 ccgcgcac--accacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 223
Qy 243 cccacacaccc--cttggacacaaagacgtgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 299
Db 224 ttcagtcgactgtgactgtgacacgaacgtcagatcgaatgtggtcgcgcgcgcgcgcgcgc 283
Qy 300 ttcacccgactgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 359
Db 284 tccacacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 343
Qy 360 tgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 413
Db 344 tgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 403
Qy 414 -----aagtggtgacacacacacacacacacacacacacacacacacacacacacacac 467
Db 404 acaaaactgagagccggagagctctcagagatggcggltgtlctcaacgacaagacgcgc 463
Qy 468 ggcgagcgctgacgctgtctgtcccggtggtgagggccgcgcgcgcgcgcgcgcgcgcgcgc 527

```

```

Db 464 acccagagatccagctgcgcaggtgtcaggaagcgcgcgcgcgcgcgcgcgcgcgcgcgc 520
Qy 528 ctgtgtgcctctgtcagatgcaagcgccctcaccgcgtctcccaacacagctcgagctcaag 587
Db 521 gtgtlctcagcgctgtcagagtgtaacacccgtctcagcaacagctcgcgcgcgcgcgcgc 580
Qy 588 gacttcggagccgagcgagcgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 647
Db 581 tttaagaagcgtgtcgcagcagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 640
Qy 648 agcgcca 654
Db 641 aaatcca 647

```

RESULT 3 US-08-468-847B-1 Sequence 1, Application US/08468847B

```

Patent No. 5780263
GENERAL INFORMATION:
: APPLICANT: Hastings, Gregg A. and Adams, Mark D.
: TITLE OF INVENTION: Human CCN-Like Growth Factor
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
: ADDRESSEE: CECCHI, STEWART & OLSTEIN
: STREET: 6 BECKER FARM ROAD
: CITY: ROSELAND
: STATE: NEW JERSEY
: COUNTRY: USA
: ZIP: 07068
COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 INCH DISKETTE
: OPERATING SYSTEM: IBM PS/2
: SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,847B
: FILING DATE: 6 June 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
ATTORNEY/AGENT INFORMATION:
: NAME: MULLINS, J.G.
: REGISTRATION NUMBER: 33,073
: REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-994-1700
: TELEFAX: 201-994-1744
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 900 BASE PAIRS
: TYPE: NUCLEIC ACID
: STRANDEDNESS: SINGLE
: TOPOLOGY: LINEAR
: MOLECULE TYPE: cDNA
US-08-468-847B-1

```

Query Match 3.0%; Score 69.6; DB 1; Length 900;

Best Local Similarity 50.0%; Pred. No. 2.5e-08; Mismatches 224; Indels 15; Gaps 2;

```

Db 121 ggtgtcagcgcttcaagaatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db 176 GCTGTGTCCTTTTAAATAATGATGCGCACAGAAATCCTTATTACATGTGGTTAAACCTG 235
Qy 181 agcctccacgcgagctgtgagacacacacacacacacacacacacacacacacacacacacac 240
Db 236 TTTCAGCAGCACCCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 295
Qy 241 ctcccaacacaccccttggagacaaagagctgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:11 ; Search time 5549.44 Seconds
(without alignments)
3919.497 Million cell updates/sec

Title: US-09-668-021-5
Perfect score: 2301
Sequence: 1 agagcctgtctactgtgaag.....caatgaatcatgaccgaaag 2301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Marched: 1022815 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est148:*
118: gb_est149:*
119: gb_est150:*
120: gb_est151:*
121: gb_est152:*
122: gb_est153:*
123: gb_est154:*
124: gb_est155:*
125: gb_est156:*
126: gb_est157:*
127: gb_est158:*
128: gb_est159:*
129: gb_est160:*
130: gb_est161:*
131: gb_est162:*
132: gb_est163:*
133: gb_est164:*
134: gb_est165:*
135: gb_est166:*
136: gb_est167:*
137: gb_est168:*
138: gb_est169:*
139: gb_est170:*
140: gb_est171:*
141: gb_est172:*
142: gb_est173:*
143: gb_est174:*
144: gb_est175:*
145: gb_est176:*
146: gb_est177:*
147: gb_est178:*
148: gb_est179:*
149: gb_est180:*
150: gb_est181:*
151: gb_est182:*
152: gb_est183:*
153: gb_est184:*
154: gb_est185:*
155: gb_est186:*
156: gb_est187:*
157: gb_est188:*
158: gb_est189:*
159: gb_est190:*
160: gb_est191:*
161: gb_est192:*
162: gb_est193:*
163: gb_est194:*
164: gb_est195:*
165: gb_est196:*
166: gb_est197:*
167: gb_est198:*
168: gb_est199:*
169: gb_est200:*
170: gb_est201:*
171: gb_est202:*
172: gb_est203:*
173: gb_est204:*
174: gb_est205:*
175: gb_est206:*
176: gb_est207:*
177: gb_est208:*
178: gb_est209:*
179: gb_est210:*
180: gb_est211:*
181: gb_est212:*
182: gb_est213:*
183: gb_est214:*
184: gb_est215:*
185: gb_est216:*
186: gb_est217:*
187: gb_est218:*
188: gb_est219:*
189: gb_est220:*

190: gb_est110:*
191: gb_est111:*
192: gb_hlc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pin1:*
208: em_gss_pin2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Human TGF- β 1
Bovine DAN/Cerberus
Marine TGF- β 1
Human secreted pro
Oligonucleotide D1
Oligonucleotide D1
Oligonucleotide D2
Oligonucleotide D2
Oligonucleotide D1
Oligonucleotide D1
Oligonucleotide D1
Oligonucleotide D2
Oligonucleotide D2
Oligonucleotide D1
Skin cell cDNA, SE
cDNA encoding murti
Skin cell cDNA, SE
Skin cell cDNA, SE
cDNA encoding small
Human small CCN-1
cDNA sequence of h
DNA encoding human
Human adult retina
Homo sapiens fetal
Human adiposin A1
Novel human polyun
Oligonucleotide D1
Human adenosine A1
Oligonucleotide D1
SAB virus gb glyco

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:52:07 ; Search time 438.73 Seconds
(without alignments)
913.093 Million cell updates/sec

Title: US-09-668-021-11

Perfect score: 638

Sequence: 1 atgcagccctactactagcccc.....ctggagaagcctactagag 638

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

rchd: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_0601:*

1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	638	100.0	638	21	AAA29058
2	556	87.1	674	21	AAA29059
3	503.2	78.9	642	21	AAA29057
4	502	78.7	2301	21	AAA29055
5	502	78.7	2301	21	AAA29056
6	502	78.7	2329	22	AAA91023
7	500.4	78.4	2301	21	AAA29061
8	500.4	78.4	2301	21	AAA29062
9	500	78.4	642	21	AAA94051
10	427	66.9	35828	21	AAA29063
11	414.2	64.9	532	21	AAA29060

12	367.4	57.6	5680	21	AAA94049	Human DAN/Cerberus
13	367.4	57.6	9301	21	AAA29064	Human TGF-beta bin
14	357.6	56.1	1104	21	AAA94050	Human DAN/Cerberus
15	106	16.6	196	21	AAC05741	Human secreted pro
16	106	16.6	198	20	AAK40842	Human secreted pro
17	83.2	13.0	954	22	AAC99782	SKN cell cDNA, SE
18	83.2	13.0	962	21	AAZ61645	cDNA encoding mur
19	83.2	13.0	962	21	AAZ61736	cDNA encoding mur
20	83.2	13.0	962	22	AAC99578	SKN cell cDNA, SE
21	83.2	13.0	962	22	AAC99569	SKN cell cDNA, SE
22	67.2	10.5	900	18	AAT47661	cDNA encoding sma
23	67.2	10.5	900	19	AAV37735	Human small CN-11
24	67.2	10.5	1719	21	AAA04698	Human adult retina
25	67.2	10.5	1774	20	AAK07571	Homo sapiens fetal
26	65.6	10.3	1692	21	AAC79467	cDNA sequence of h
27	65.6	10.3	1710	20	AAK84238	cDNA encoding human
28	57	8.9	412	22	AAF66218	Novel human polyu
29	52.4	8.2	71989	21	AAA29349	S. cellulosus DNA
30	47.8	7.5	37856	21	AAA11992	Platenolide syntha
31	47.2	7.4	44377	18	AAT78508	Platenolide syntha
32	47.2	7.4	44377	18	AAT80414	Human adenosine A1
33	46.4	7.3	114955	20	AAK53491	dx gene of Therm
34	46.2	7.2	2007	19	AAV71089	Thermus thermophil
35	46.2	7.2	2007	20	AAZ30905	Partial human Notc
36	44.4	7.0	5617	19	AAV57163	Human Notch3 cDNA
37	44.4	7.0	8091	19	AAV57001	Human KILMP cDNA
38	44.2	6.9	3930	21	AAZ44744	Monkey MANGO 245 D
39	43.8	6.9	1388	22	AAF28262	Nucleotide sequenc
40	43.2	6.8	58857	21	AAK58471	cDNA encoding oste
41	43	6.7	2075	18	AAV75444	Orf virus genomic
42	43	6.7	4020	18	AAI91361	Human adenosine A1
43	43	6.7	114955	20	AAK53491	S. venezuelae mact
44	42.8	6.7	11220	21	AAZ87298	
45	42.8	6.7	36778	21	AAZ87318	

ALIGNMENTS

RESULT 1	
AAA29058	
ID AAA29058 standard; cDNA; 638 BP.	
XX	
AC AAA29058;	
XX	
DT 12-SEP-2000 (first entry)	
XX	
DE Murine TGF-beta binding protein (BEER) cDNA.	
XX	
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization; ss.	
XX	
OS Mus musculus.	
XX	
FH key	location/Qualifiers
FT CDS	1..636
FT	/*tag= a
FT	/product= TGF-beta_binding_protein
XX	
PN WO200032773-A1.	
XX	
PD 08-JUN-2000.	
XX	
PD 24-NOV-1999;	99WO-US27990.
XX	
PR 27-NOV-1998;	98US-0110283.
XX	
PA (DARW-) DARWIN DISCOVERY LTD.	
XX	
FI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
PI Van Ness J, Winkler DG;	
XX	
DR WPI; 2000-412321/35.	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:52:21 ; Search time 438.73 Seconds
(without alignments)
964.615 Million cell updates/sec

Title: US-09-668-021-13

Perfect score: 674
Sequence: 1 gagaccgagtgccctcct.....agctggagaaagcctactag 674

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

riched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_0601.*
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	674	100.0	674	21	AAA29059 Rat TGF-beta bind1
2	556	82.5	638	21	AAA29058 Murine TGF-beta b1
3	536.2	79.6	2301	21	AAA29055 Human TGF-beta bin
4	536.2	79.6	2329	22	AAA91023 Human secreted pro
5	534.6	79.3	2301	21	AAA29061 Mutant human TGF-b
6	534.6	79.3	2301	21	AAA29062 Human TGF-beta bin
7	533	79.1	2301	21	AAA29056 Human TGF-beta bin
8	530	78.6	642	21	AAA29057 Vervet TGF-beta b1
9	525.2	77.9	642	21	AAA94051 Human DAN/Cerberus
10	424.6	63.0	532	21	AAA29060 Bovine TGF-beta b1
11	407.4	60.4	35828	21	AAA29063 Murine TGF-beta b1

12	372.6	55.3	5680	21	AAA94049 Human DAN/Cerberus
13	372.6	55.3	9301	21	AAA29064 Human TGF-beta bin
14	364.4	54.1	1104	21	AAA94050 Human DAN/Cerberus
15	119	17.7	196	21	AAAC05741 Human secreted pro
16	119	17.7	198	20	AAAX0842 Human secreted pro
17	82.8	12.3	954	22	AAAC99782 Skin cell CDNA, SE
18	82.8	12.3	962	21	AAZ61645 CDNA encoding mur1
19	82.8	12.3	962	21	AAZ61736 CDNA encoding mur1
20	82.8	12.3	962	22	AAAC9578 Skin cell CDNA, SE
21	82.8	12.3	962	22	AAAC9669 Skin cell CDNA, SE
22	64	9.5	900	18	AAAT47661 CDNA encoding sma1
23	64	9.5	1719	21	AAV37735 Human small CCN-1i
24	64	9.5	1719	21	AAA40498 Human adult fetina
25	64	9.5	1719	20	AAAX07571 Homo sapiens fetal
26	63.2	9.4	1692	21	AAAC79467 CDNA sequence of h
27	63.2	9.4	1710	20	AAAC84238 DNA encoding human
28	56.2	8.3	412	22	AAAF66218 Novel human polynu
29	54.2	8.0	71989	21	AAA29349 Sorangium cellulos
30	54	8.0	37856	21	AAA11992 S. cellulosum DNA
31	49.2	7.3	5617	19	AAV57163 Partial human Notc
32	49.2	7.3	8091	19	AAV57001 Human Notch3 CDNA
33	47.2	7.0	1125	22	AAAC84259 Signal transductio
34	47	7.0	1380	19	AAV67187 M. luteus salt-res
35	46.6	6.9	674	21	AAA29059 Rat TGF-beta bind1
36	46.6	6.9	1525	21	AAAC44071 Zee mays DNA fragm
37	46.2	6.9	2007	19	AAV71085 dnax gene of Therm
38	46.2	6.9	2007	20	AAZ30905 Thermus thermophil
39	46	6.8	11220	21	AAZ87298 S. venezuelae macr
40	46	6.8	36778	21	AAZ87318 S. venezuelae pik
41	46	6.8	37948	21	AAZ87285 S. venezuelae pik
42	46	6.8	38506	21	AAZ75633 Nucleotide sequenc
43	46	6.8	38506	21	AAZ56001 Recombinant cosmid
44	45.4	6.7	18660	21	AAZ58472 Nucleotide sequenc
45	45.4	6.7	44377	18	AAZ78508 Platenolide syntha

ALIGNMENTS

RESULT 1					
ID	AAA29059	standard;	CDNA:	674	BP.
AC	AAA29059;				
DF	12-SEP-2000	(first entry)			
DE	Rat TGF-beta binding protein (BEER) CDNA.				
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;				
KM	BEER, gene therapy; antisense therapy; fracture; bone mineralization; ss.				
XX					
OS	Rattus norvegicus.				
FH	Key	Location/Qualifiers			
FT	CDS	33..674			
FT		/*tag= a			
XX		/product= TGF-beta_binding_protein			
PN	MO200032773-A1.				
XX					
PD	08-JUN-2000.				
XX					
PF	24-NOV-1999;	99WO-US27990.			
XX					
PR	27-NOV-1998;	98US-0110283.			
XX					
PA	(DARW-) DARWIN DISCOVERY LTD.				
XX					
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;				
XX	Van Ness J, Winkler DG;				
DR	WPI; 2000-412321/35.				

[illegible]

```

1  RESULT 2
2  US-09-188-930-209
3  Sequence 209, Application US/09188930A
4  Patent No. 6150502
5  GENERAL INFORMATION:
6  APPLICANT: Watson, James D.
7  APPLICANT: Strachan, Lorna
8  APPLICANT: Sleeman, Matthew
9  APPLICANT: Onrust, Rene
10 TITLE OF INVENTION: Compositions Isolated From Skin Cells
11 TITLE OF INVENTION: and Methods For Their Use
12 FILE REFERENCE: 11000.1011c1
13 CURRENT APPLICATION NUMBER: US/09/188,930A
14 CURRENT FILING DATE: 1998-11-09
15 NUMBER OF SEQ ID NOS: 348
16 SOFTWARE: FastSeq for Windows Version 3.0
17 SEQ ID NO 209
18 LENGTH: 962
19 TYPE: DNA
20 ORGANISM: Mouse
21 US-09-188-930-209

```

[illegible][illegible]

```

1      RESULT      3
2      US-08-468-847B-1
3      : Sequence 1, Application US/08468847B
4      : Patent No. 5780263
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Haslings, Gregg A. and Adams, Mark D.
8      : TITLE OF INVENTION: Human CCN-Like Growth Factor
9      : NUMBER OF SEQUENCES: 20
10     : CORRESPONDENCE ADDRESS:
11     : ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
12     : STREET: 6 BECKER FARM ROAD
13     : CITY: ROSELAND
14     : STATE: NEW JERSEY
15     :
16     : COUNTRY: USA
17     : ZIP: 07068
18     :
19     : COMPUTER READABLE FORM:
20     : MEDIUM TYPE: 3.5 INCH DISKETTE
21     : COMPUTER: IBM PS/2
22     : OPERATING SYSTEM: MS-DOS
23     : SOFTWARE: WORD PERFECT 5.1
24     :
25     : CURRENT APPLICATION DATA:
26     : APPLICATION NUMBER: US/08/468,847B
27     : FILING DATE: 6 June 1995
28     : CLASSIFICATION: 435
29     : PRIOR APPLICATION DATA:
30     : APPLICATION NUMBER:
31     : FILING DATE:
32     : ATTORNEY/AGENT INFORMATION:
33     : NAME: MULLINS, J.G.
34     : REGISTRATION NUMBER: 33,073
35     : REFERENCE/DOCKET NUMBER: 325800-442
36     : TELECOMMUNICATION INFORMATION:
37     : TELEPHONE: 201-994-1700
38     : TELEFAX: 201-994-1744
39     : INFORMATION FOR SEQ ID NO: 1:
40     : SEQUENCE CHARACTERISTICS:
41     : LENGTH: 900 BASE PAIRS
42     : TYPE: NUCLEIC ACID
43     : STRANDEDNESS: SINGLE
44     : TOPOLOGY: LINEAR
45     :
46     : MOLECULE TYPE: CDNA
47     :
48     : US-08-468-847B-1

```

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:52:36 ; Search time 438.73 Seconds
(without alignments)
761.388 Million cell updates/sec

Title: US-09-668-021-15

Perfect score: 532
Sequence: 1 agaatgatccaccagaatc.....accagaagccagcggccga 532

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Archived: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_0601:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	532	100.0	532	21	AAA29060
2	467.8	87.9	642	21	AAA94051
3	467.8	87.9	2301	21	AAA29055
4	467.8	87.9	2301	21	AAA29061
5	467.8	87.9	2329	22	AAA91033
6	466.2	87.6	2301	21	AAA29056
7	466.2	87.6	2301	21	AAA29052
8	461.4	86.7	642	21	AAA29057
9	461.4	86.7	674	21	AAA29059
10	414.2	77.9	638	21	AAA29058
11	375.2	70.5	5680	21	AAA94049

12	375.2	70.5	9301	21	AAA29064	Human TGF-beta bin
13	364.6	68.5	1104	21	AAA94050	Human DAN/Cerberus
14	332.4	62.5	35828	21	AAA29063	Murine TGF-beta bl
15	93	17.5	954	22	AAC99782	SKIN cell cDNA, SE
16	93	17.5	962	21	AA61645	cDNA encoding murt
17	93	17.5	962	21	AA61736	cDNA encoding murt
18	93	17.5	962	22	AAC99578	SKIN cell cDNA, SE
19	93	17.5	962	22	AAC99669	SKIN cell cDNA, SE
20	72.2	13.6	900	18	AA147661	cDNA encoding sma1
21	72.2	13.6	900	19	AAV37735	Human small CCN-11
22	72.2	13.6	1692	21	AAC79467	cDNA sequence of h
23	72.2	13.6	1710	20	AA584238	Human adult retina
24	72.2	13.6	1719	21	AAA40498	Human adult retina
25	72.2	13.6	1774	20	AA507571	Homo sapiens fetal
26	61	11.5	412	22	AA66218	Novel human polyn
27	55.6	10.5	37856	21	AAA11992	S. cellulosum DNA
28	53.2	10.4	114955	20	AA53491	Human adenosine A1
29	49.6	9.3	114955	20	AA53491	Human adenosine A1
30	49.2	9.2	198	20	AA540842	Human secreted pro
31	48.8	9.2	196	21	AAC05741	Human secreted pro
32	48	9.0	19053	21	AA59146	Nucleotide sequenc
33	48	9.0	19053	21	AA59146	Sorandium cellulos
34	47.8	9.0	71989	21	AAA29349	Human secreted pro
35	46.6	8.8	1479	21	AA51609	HIV synthetic Gag
36	46.6	8.8	1479	21	AA51625	HIV codon-optimiz
37	45.2	8.5	3861	14	AA031543	Lipopolysaccharide
38	44.8	8.4	2140	15	AA063180	Alpha 1a adrenegrl
39	44.8	8.4	2140	16	AA703130	Human alpha-1a adr
40	44.8	8.4	2140	22	AAE24205	Human alpha-1a adr
41	44.4	8.3	1485	21	AA61590	DNA encoding a hum
42	44.4	8.3	1485	21	AA61590	DNA encoding a hum
43	44.4	8.3	2277	19	AAV13836	Homo sapiens mamma
44	44.4	8.3	2277	19	AAV05372	Human telomerase p
45	44.4	8.3	44377	18	AA178508	Platenolide synth

ALIGNMENTS

RESULT 1	
AAA29060	
ID	AAA29060 standard: cDNA: 532 BP.
AC	AAA29060;
XX	
DT	12-SEP-2000 (first entry)
XX	
DE	Bovine TGF-beta binding protein (BEER) cDNA.
XX	
KW	osteopathic: transforming growth factor-beta; TGF-beta: binding protein;
KW	BEER; gene therapy: antisense therapy; fracture; bone mineralization; ss.
XX	
OS	Bos taurus.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..530
FT	/*tag= a
FT	/partial
FT	/product= TGF-beta_binding-protein
XX	
PN	W0200032773-A1.
XX	
PD	08-JUN-2000.
XX	
XX	24-NOV-1999;
PF	99WO-US27990.
XX	
XX	27-NOV-1998;
PR	98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
XX	
PI	Brunkow ME, Gals DJ, Kovacevich B, Mulligan JT, Paepfer BW;
PI	Van Ness J, Winkler DG;
XX	

DR WPI: 2000-412321/35.
DR P-PSDB; AAT96434.

PT Nucleic acids (1) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteopenia, osteoporosis and fractures

PS Claim 1: Page 126; 162pp; English.

CC This cDNA encodes a bovine transforming growth factor-beta (TGF-beta) binding protein designated bBER. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate bBER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of bBER by CC expressing inactive proteins or to supplement the patient's own production of bBER polypeptides. The nucleic acids may be used for recombinant CC production of bBER, gene therapy, antisense therapy, as probes for CC diagnostic assays and for functional studies. bBER may be used to raise CC antibodies and for identification of bBER modulators. bBER antagonists CC may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

SO Sequence 532 BP; 95 A; 199 C; 176 G; 62 T; 0 other;

Query Match 100.0%; Score 532; DB 21; Length 532;
Best Local Similarity 100.0%; Pred. No. 4.8e-95;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaatgagtcacagaatcatcccgagctgagagctaccgagccctcgccagc 60
DB 1 agaatgagtcacagaatcatcccgagctgagagctaccgagccctcgccagc 60
QY 61 tgaacaaagaagacatgaaacggcgaggaagagagagacccaccacccttgg 120
DB 61 tgaacaaagaagacatgaaacggcgaggaagagagagacccaccacccttgg 120
QY 121 agaccaaagagcctcgcagctacagctccgagagctcacttaccgcgtacgtaccg 180
DB 121 agaccaaagagcctcgcagctacagctccgagagctcacttaccgcgtacgtaccg 180
QY 121 agaccaaagagcctcgcagctacagctccgagagctcacttaccgcgtacgtaccg 180
DB 121 agaccaaagagcctcgcagctacagctccgagagctcacttaccgcgtacgtaccg 180
QY 181 atgggacgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
DB 181 atgggacgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
QY 241 cggcgagcctgctgcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
DB 241 cggcgagcctgctgcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
QY 301 acttcgctgcatcccgagcagcagcagcagcagcagcagcagcagcagcagcagc 360
DB 301 acttcgctgcatcccgagcagcagcagcagcagcagcagcagcagcagcagcagc 360
QY 361 ggcgcgagc 420
DB 361 ggcgcgagc 420
QY 421 ctgccttcacacacagctcagctcagctcagctcagctcagctcagctcagctcag 480
DB 421 ctgccttcacacacagctcagctcagctcagctcagctcagctcagctcagctcag 480
QY 481 cggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 532
DB 481 cggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 532

RESULT 2
AAA94051
ID AAA94051 standard; DNA; 642 BP.
XX

AC AAA94051;
XX 15-JAN-2001 (first entry)
DT Human DAN/Cerberus-related protein 6 (hDcr6) cDNA exons 1 and 4.
XX
DE Human DAN/Cerberus-related protein 6 (hDcr6) cDNA exons 1 and 4.
XX
KW Human: DNA/Cerberus-related protein 6; hDcr6; morphogenic protein;
KW antagonist; BMP; cell growth; cell differentiation; bone formation;
KW gene therapy; ss.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..642
FT /tag= a
FT /product= "hDcr6 #2"
FT /partial
XX
XX W020005193-A2.
XX
XX 21-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-US05537.
XX
XX 12-MAR-1999; 99US-0124116.
XX
XX (REG-) REGENERON PHARM INC.
XX
XX Economides AN;
XX
XX WPI: 2000-638179/61.
XX
XX P-PSDB; AAB26106.
XX
XX Novel isolated, human DNA/Cerberus related protein 6 which include
PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
PT acids encoding the proteins which are useful as probes and primers
XX
XX Claim 2; Fig 3; 40pp; English.
XX
XX The present sequence comprises exons 1 and 4 of the human
CC DAN/Cerberus-related protein 6 (hDcr6) coding sequence. It was isolated
CC from a human kidney cDNA library containing exons 1 and 4 of the
CC sequence. hDcr6 is closely related to the DAN and DCR5 proteins, both of
CC which act as antagonists of morphogenic proteins such as BMP. It is
CC possible that the hDcr6 gene and protein can be used as immunogens,
CC modulators of cell function, growth and differentiation, to reduce
CC undesirable bone formation, to identify DCR6 binding agents, in
CC diagnosis, and in gene therapy.
XX
XX Sequence 642 BP; 115 A; 230 C; 211 G; 86 T; 0 other;

Query Match 87.9%; Score 467.8; DB 21; Length 642;
Best Local Similarity 93.5%; Pred. No. 1.3e-82;
Matches 500; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

QY 1 agaatgagtcacagaatcatcccgagcctgagagctaccgagcctctgcagagc 60
DB 1 agaatgagtcacagaatcatcccgagcctgagagctaccgagcctctgcagagc 60
QY 61 t---gaacaaagaagacatgaaacggcgaggaacgaggaagagcctccaccaccct 117
DB 61 t---gaacaaagaagacatgaaacggcgaggaacgaggaagagcctccaccaccct 117
QY 149 tgaagaacaaagaagacatgaaacggcgaggaacgaggaagagcctccaccaccct 208
DB 149 tgaagaacaaagaagacatgaaacggcgaggaacgaggaagagcctccaccaccct 208
QY 178 cgaatggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 237
DB 178 cgaatggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 237
QY 269 cgaatggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 328
DB 269 cgaatggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 328
QY 238 gcccgagcagcctgctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 297
DB 238 gcccgagcagcctgctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 297

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 01:30:44 ; Search time 222.63 Seconds
(without alignments)
541.195 Million cell updates/sec

Title: US-09-668-021-15

Perfect score: 532

Sequence: 1 agaataatgacccagcaaatc.....accaaagccagcgccgca 532

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Archived: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	17.5	962	US-09-188-930-40	Sequence 40, Appl
2	93	17.5	962	US-09-188-930-209	Sequence 209, App
3	72.2	13.6	900	US-08-468-847B-1	Sequence 1, Appl
4	48	9.0	71989	US-09-443-501A-2	Sequence 2, Appl
5	47.8	9.0	71989	US-09-443-501A-2	Sequence 2, Appl
6	45.2	8.5	2150	US-08-318-837-1	Sequence 1, Appl
7	44.8	8.4	2140	US-08-318-837-1	Sequence 1, Appl
8	44.8	8.4	2140	US-08-228-932-1	Sequence 1, Appl
9	44.8	8.4	2140	US-08-468-939-1	Sequence 1, Appl
10	44.8	8.4	2140	US-08-406-855A-1	Sequence 1, Appl
11	44.8	8.4	2140	US-08-722-190-1	Sequence 1, Appl
12	44.8	8.4	2140	US-08-244-354-1	Sequence 1, Appl
13	44.8	8.4	2140	US-09-206-899-1	Sequence 1, Appl
14	44.8	8.4	2140	PCT-US95-04203-1	Sequence 1, Appl
15	44.4	8.3	2277	US-08-676-967-5	Sequence 5, Appl
16	44.4	8.3	2277	US-08-676-974-5	Sequence 5, Appl
17	44.4	8.3	2277	US-09-098-487-5	Sequence 5, Appl
18	44.4	8.3	44377	US-08-804-227C-7	Sequence 7, Appl
19	44.4	8.3	44377	US-08-804-198-1	Sequence 7, Appl
20	44.2	8.3	835	US-08-417-822A-1	Sequence 1, Appl
21	44.2	8.3	11907	US-08-061-376-4	Sequence 4, Appl
22	43.8	8.2	43280	US-08-804-227C-1	Sequence 1, Appl
23	43.4	8.2	13842	US-09-105-537-30	Sequence 30, Appl
24	43.4	8.2	36778	US-09-105-537-5	Sequence 5, Appl
25	43.4	8.2	38506	US-09-320-878-19	Sequence 19, Appl
26	43.2	8.1	1780	US-08-933-821-5	Sequence 5, Appl
27	43.2	8.1	1780	US-08-960-507-5	Sequence 5, Appl

28	43.2	8.1	2849	US-08-809-286B-1	Sequence 1, Appl
29	43.2	8.1	2943	US-08-042-747A-7	Sequence 7, Appl
30	43.2	8.1	3468	US-07-951-715A-2	Sequence 2, Appl
31	43.2	8.1	3468	US-08-459-448A-2	Sequence 2, Appl
32	43.2	8.1	3468	US-08-459-595A-2	Sequence 2, Appl
33	43.2	8.1	3468	US-08-459-504B-2	Sequence 2, Appl
34	43.2	8.1	3468	US-08-459-444-2	Sequence 2, Appl
35	43.2	8.1	3468	US-09-053-549-3	Sequence 3, Appl
36	43.2	8.1	11958	US-09-134-246-8	Sequence 8, Appl
37	42.8	8.0	1738	US-08-334-698-3	Sequence 3, Appl
38	42.8	8.0	1738	US-08-228-932-3	Sequence 3, Appl
39	42.8	8.0	1738	US-08-468-939-3	Sequence 3, Appl
40	42.8	8.0	1738	US-08-406-855A-3	Sequence 3, Appl
41	42.8	8.0	1738	US-08-722-190-3	Sequence 3, Appl
42	42.8	8.0	1738	US-08-244-354-3	Sequence 3, Appl
43	42.8	8.0	1738	US-09-206-899-3	Sequence 3, Appl
44	42.8	8.0	1738	PCT-US95-04203-3	Sequence 3, Appl
45	42.6	8.0	1820	US-08-173-508-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1	US-09-188-930-40	Sequence 40, Application US/09188930A
Patent No. 6150502		
GENERAL INFORMATION:		
APPLICANT: Watson, James D.		
APPLICANT: Strachan, Lorna		
APPLICANT: Sleeman, Matthew		
APPLICANT: Onrust, Rene		
APPLICANT: Marison, James Greg		
TITLE OF INVENTION: Compositions Isolated From Skin Cells		
FILE REFERENCE: 11000.1011c1		
CURRENT APPLICATION NUMBER: US/09/188,930A		
CURRENT FILING DATE: 1998-11-09		
NUMBER OF SEQ ID NOS: 348		
SOFTWARE: FastSeq for Windows Version 3.0		
SEQ ID NO 40		
LENGTH: 962		
TYPE: DNA		
ORGANISM: mouse		
US-09-188-930-40		
Query Match	17.5%	Score 93; DB 3; Length 962;
Best Local Similarity	56.6%	Pred. No. 2e-11;
Matches 263; Conservative	0;	Mismatches 180; Indels 22; Gaps 4;
QY	1	agaataatgacccagcaaatcctcccgagctgagcgagctcccgagctctgcagagc 60
DB	120	aaataatgacccagcaaatcctcccgagctgagcgagctcccgagctctgcagagc 179
QY	61	tgaacacacagacacatgaacacacacacacacacacacacacacacacacacac 120
DB	180	gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 237
QY	121	agac 175
DB	238	actgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 297
QY	176	gacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 235
DB	298	ttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 357
QY	236	cagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 283
DB	358	ctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 417
QY	284	gcgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 343

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:46:18 ; Search time 54.26 Seconds
(without alignments)
299.027 Million cell updates/sec

Title: US-09-668-021-2

Perfect score: 1167
Sequence: 1 MQLPLACLVCLVHTAFRV.....KPRPRARSAKANQAELENAY 213

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

rchd: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270	23.1	134	2 T08710	hypothetical prote
2	94.5	8.1	159	2 I51373	luteinizing hormon
3	87.5	7.5	158	2 A61091	lutropin beta chai
4	87.5	7.5	601	2 T22025	hypothetical prote
5	86.5	7.4	1138	2 G85077	probable polypept
6	85.5	7.3	715	2 T12534	hypothetical prote
7	85	7.3	1469	2 B36655	slit protein 2 pre
8	83.5	7.2	824	2 T10020	leucine--trna 11ga
9	83.5	7.2	972	2 T10023	slit protein 1 pre
10	82.5	7.1	1480	2 A36655	hypothetical prote
11	82	7.0	2531	2 T16743	maturing type A prot
12	81.5	7.0	520	2 S47142	hypothetical prote
13	81.5	7.0	969	2 A70912	probable leus prot
14	81	6.9	270	2 S71793	head-inducing fact
15	81	6.9	839	2 I50590	class I INCEMP pro
16	80	6.9	2333	1 GNNY2T	genome polypeptin
17	79	6.8	118	2 PNO141	lutropin beta chai
18	79	6.8	178	2 I51824	tumor-suppressive
19	79	6.8	178	2 A47291	probable finger pr
20	79	6.8	866	2 T29197	hypothetical prote
21	78.5	6.7	166	2 I51242	luteinizing hormon
22	78.5	6.7	690	2 T33321	hypothetical prote
23	78.5	6.7	1122	2 T14180	exit protein - Myc
24	78	6.7	877	2 I50591	class II INCEMP pr
25	78	6.7	1142	2 T30272	hypothetical prote
26	78	6.7	1210	1 GCHUE	epidermal growth f
27	77.5	6.6	1042	2 A57534	mutin 5AC (clone L
28	77	6.6	118	2 PNO139	lutropin beta chai
29	77	6.6	343	2 G35070	apolipoprotein H-r

30	77	6.6	510	2 A42750	insulinoma-associa
31	77	6.6	562	1 UKHUT	t-plasminogen acti
32	77	6.6	702	2 E72775	probable helicase
33	77	6.6	864	2 H85335	hypothetical prote
34	77	6.6	864	2 T04518	hypothetical prote
35	76.5	6.6	402	2 T48729	hypothetical prote
36	76.5	6.6	1223	2 S29717	adenylate cyclase
37	76	6.5	398	2 A35281	integumentary muc1
38	76	6.5	2142	2 B35098	MHC class III hist
39	75.5	6.5	403	2 C70832	hypothetical prote
40	75.5	6.5	646	2 T02398	hypothetical prote
41	75.5	6.5	1436	2 A46496	antigen wC1.1 prec
42	75	6.4	992	2 A31666	hypothetical prote
43	75	6.4	2153	2 T30074	hypothetical prote
44	74.5	6.4	1063	2 T03743	bifocal protein -
45	74	6.3	269	2 E69381	hypothetical prote

ALIGNMENTS

RESULT 1
T08710
hypothetical protein DKFP564D206.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #extl_change 13-Aug-1999
C:Accession: T08710
R:Wambull, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
Submitted to the Protein Sequence Database, March 1999
A:Reference number: Z16471
A:Accession: T08710
A:Molecule type: mRNA
A:Residues: 1-134 <NAME>
A:Cross-references: EMBL:AL050024
A:Experimental source: Fetal brain; clone DKFP564D206
C:Genetics:
A>Note: DKFP564D206.1

Query Match 23.1% Score 270. DB 2; Length 134;
Best Local Similarity 44.2% Pred. No. 5.2e-17;
Matches 57; Conservative 23; Mismatches 43; Indels 6; Gaps 4;

QY 80 CRELHFTRYVDGPRRSKAVTELVCSGCGPARLLPNAIGRG---KMW-RPSGDPFRCI 135
||| |::|| | | |::||:| | |::|| | | |::|| | | |::|| | | |::||
Db 3 CRELSTRTKITSDCGCTISPLKELVCAEGELLPLVLPNWTGGYGTIKYRSRRSSQEMRCY 62
|: |::|| | | | |::|| | | |::|| | | |::|| | | |::|| | | |::||
QY 136 PDRTAARVOLLCPGEAPPARAVRYASCKRRLTFPHNQSLEKDGTEAARPKGRK 195
|: |::|| | | | |::|| | | |::|| | | |::|| | | |::|| | | |::||
Db 63 NDKRTTORIOLQCGDG-STRTYKITVTYACKCKRYTROHNESSHNFSMPAKPVQHRE 121
QY 196 RPRR-RSAK 203
| |::|| | | | |::|| | | | |::|| | | | |::|| | | | |::||
Db 122 RKRASKSK 130

RESULT 2
I51373
luteinizing hormone beta subunit - turkey
C:Species: Melalegris gallipavo (common turkey)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #extl_change 16-Jul-1999
C:Accession: I51373
R:You, S.; Foster, L.K.; Silsby, J.L.; el Halawani, M.E.; Foster, D.N.
J. Mol. Endocrinol. 14, 117-129, 1995
A:Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonado
A:Reference number: I51373; MID:95290073
A:Accession: I51373
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-159 <YOU>
A:Cross-references: GB:U35519; NID:9530952; PIDN:AAA74125.1; PID:9530953
C:Genetics:
A:Gene: LH-beta

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:48 ; Search time 31.13 Seconds

(without alignments)
234.385 Million cell updates/sec

Title: US-09-668-021-2
Perfect score: 1167
Sequence: 1 MOLPLALCVLIVHTAFRV.....KPRPARSAKANOAELENNY 213

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94.5	8.1	159	LSHB_MELGA	P45646 meleagris g
2	87	7.5	958	HIG_DROME	Q09101 drosophila
3	85	7.3	180	DAN_HUMAN	P41271 homo sapien
4	83.5	7.2	141	LSHB_PRTIV	O46482 trichosurus
5	83.5	7.2	972	SVL_MYCLE	O50192 mycobacteri
6	82.5	7.1	1480	SLIT_DROME	P24014 drosophila
7	81.5	7.0	870	BCAL_HUMAN	P56945 homo sapien
8	81.5	7.0	969	SVL_MYCTU	P71698 mycobacteri
9	80	6.9	2333	POLG_FMDV1	P03306 f genome po
10	79.5	6.8	318	HR23_MOUSE	P42582 mus musculu
11	79	6.8	118	LSHB_PHYCA	P23330 physeter ca
12	79	6.8	178	DAN_RAT	O06880 rattus norv
13	79	6.8	355	KLF2_HUMAN	O95671 homo sapien
14	78.5	6.7	166	LSHB_COTUA	P45657 coturnix co
15	78	6.7	187	INCE_CHICK	P53352 gallus gall
16	78	6.7	1210	SEPR_HUMAN	P00533 homo sapien
17	77.5	6.6	951	SEPR_HUMAN	Q12872 homo sapien
18	77	6.6	118	LSHB_BALAC	P33086 balaenopter
19	77	6.6	510	ITAI_HUMAN	P00110 homo sapien
20	77	6.6	562	TPA_HUMAN	P00750 homo sapien
21	76	6.5	398	MUB1_XENLA	P38565 xenopus lae
22	76	6.5	2142	BAT2_HUMAN	P48634 homo sapien
23	75.5	6.5	318	HK25_RAT	O35767 rattus norv
24	75.5	6.5	354	ATH1_HUMAN	O92854 homo sapien
25	75.5	6.5	1436	WC11_BOVIN	P30205 bos taurus
26	75	6.4	141	LSHB_CERST	O77833 ceratotheri
27	75	6.4	992	EBN6_EBV	P03204 Epstein-Bar
28	74	6.3	138	LSHB_MACRO	O46483 macropus ru
29	74	6.3	470	KLF4_HUMAN	O43474 homo sapien
30	74	6.3	474	KLF4_MOUSE	O60793 mus musculu
31	74	6.3	544	H15_DROME	O94890 drosophila
32	74	6.3	2715	TRX2_HUMAN	O94890 homo sapien
33	73.5	6.3	375	CE10_CHICK	P19336 gallus gall

34	73.5	6.3	682	1	SNK_MOUSE	P53351 mus musculu
35	73.5	6.3	682	1	SNK_RAT	O94012 rattus norv
36	73.5	6.3	685	1	SNK_HUMAN	O94973 homo sapien
37	73	6.3	178	1	DAN_MOUSE	O61477 mus musculu
38	73	6.3	191	1	CBX5_HUMAN	P45973 homo sapien
39	73	6.3	301	1	UL49_HSV11	P10233 herpes simp
40	73	6.3	2336	1	POLG_FMDV2	P49303 f genome po
41	72.5	6.2	128	1	LSHB_STRCA	P80664 struthio ca
42	72.5	6.2	507	1	DAF_CAVPO	O60401 cavia porce
43	72.5	6.2	543	1	PKA_STRCO	P54739 streptomyce
44	72	6.2	243	1	TOMB_PSEPU	O05613 pseudomonas
45	72	6.2	379	1	CYRE_MOUSE	P18406 mus musculu

ALIGNMENTS

RESULT 1	LSHB_MELGA	STANDARD;	PRT;	159 AA.	
ID	LSHB_MELGA				
AC	P45646;				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH- BETA) (LSH-B) (LH-B).				
GN	LHB.				
OS	Meleagris gallipavo (Common turkey).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.				
OX	NCBI_TaxID=9103;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
KC	TISUE=Pituitary;				
RX	MEDLINE=95290073; PubMed=7772235;				
RA	You S., Foster L.K., Sillsby J.L., el Halaoui M.E., Foster D.N.;				
RT	*Sequence analysis of the turkey LH beta subunit and its regulation by gonadotrophin-releasing hormone and prolactin in cultured pituitary cells.*;				
RT	J. Mol. Endocrinol. 14:117-129(1995).				
RL	- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.				
CC	- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.				
CC	- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; L35519; AA74125.1; ALT_INIT.				
DR	HSSP; F01233; IHRP.				
DR	InterPro; IPR000359; -.				
DR	InterPro; IPR001545; -.				
DR	InterPro; IPR002400; -.				
DR	Pfam; PF00007; Cys_knot; 1.				
DR	PRINTS; PR00438; GRCYSKNOR.				
DR	PROSITE; PS00681; GLYCO_HORMONE_BETA_1; 1.				
DR	PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.				
KW	Hormone; Signal; Glycoprotein.				
KW	SIGNAL				
FT	CHAIN	1	39		POTENTIAL.
FT	DISULFID	40	159		LUTROPIN BETA CHAIN.
FT	DISULFID	48	96		BY SIMILARITY.
FT	DISULFID	62	111		BY SIMILARITY.
FT	DISULFID	65	149		BY SIMILARITY.
FT	DISULFID	73	127		BY SIMILARITY.
FT	DISULFID	77	129		BY SIMILARITY.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:47:23 (Search time 89.34 Seconds
315.435 Million cell updates/sec)

Title: US-09-668-021-2
Perfect score: 1167
Sequence: 1 MOLPLALICVLCVLTAFRV.....KPPRRASAKANQAELENNAY 213

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organellar:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_unclassified:*
14: sp_vertebrate:*
15: sp_virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	270	23.1	134	4	Q9Y3U3	Q9Y3U3 homo sapien
2	106	9.1	272	13	Q9PWB0	Q9PWB0 gallus gall
3	104	8.9	272	13	Q9PWB0	Q9PWB0 gallus gall
4	103	8.8	168	4	Q9H772	Q9H772 homo sapien
5	99.5	8.5	493	4	Q9URM9	Q9URM9 homo sapien
6	98	8.4	168	11	Q88273	Q88273 mus musculu
7	96	8.2	184	4	Q60565	Q60565 homo sapien
8	95.5	8.2	493	4	Q9S376	Q9S376 homo sapien
9	94.5	8.1	481	4	Q9H826	Q9H826 homo sapien
10	91	7.8	215	8	Q9MFB6	Q9MFB6 beta vulgar
11	90	7.8	1114	11	Q9JRW7	Q9JRW7 mus musculu
12	90	7.7	184	11	Q85793	Q85793 mus musculu
13	90	7.7	184	11	Q70326	Q70326 mus musculu
14	89.5	7.7	492	11	Q921K6	Q921K6 mus musculu
15	89	7.6	184	13	Q73755	Q73755 gallus gall
16	86.5	7.4	182	13	Q73754	Q73754 xenopus lae
17	86.5	7.4	1138	10	Q92PF9	Q92PF9 arabidopsis
18	85.5	7.3	715	4	Q9Y4O3	Q9Y4O3 homo sapien
19	83	7.1	958	5	Q9V560	Q9V560 drosophila

20	82.5	7.1	329	2	Q9WY60	Q9WY60 acetobacter
21	82.5	7.1	1480	5	Q9V7F8	Q9V7F8 drosophila
22	82.5	7.1	1504	5	Q9XIV4	Q9XIV4 drosophila
23	82.5	7.1	1504	5	Q9V7F9	Q9V7F9 drosophila
24	82	7.0	267	4	Q95813	Q95813 homo sapien
25	82	7.0	988	6	Q97867	Q97867 sus scrofa
26	82	7.0	5231	5	Q21980	Q21980 caenorhabdi
27	81.5	7.0	520	3	Q00352	Q00352 coprinus ci
28	81.5	7.0	661	5	Q9V7U8	Q9V7U8 drosophila
29	81	6.9	270	13	P70041	P70041 xenopus lae
30	81	6.9	335	4	Q9UKR6	Q9UKR6 homo sapien
31	81	6.9	1096	10	Q9XE24	Q9XE24 oryza sativ
32	80	6.9	712	4	Q9Y557	Q9Y557 homo sapien
33	80	6.9	720	4	Q9NYZ3	Q9NYZ3 homo sapien
34	80	6.9	1440	5	Q20204	Q20204 caenorhabdi
35	79.5	6.8	500	10	Q9M6C2	Q9M6C2 trifolium r
36	79.5	6.8	543	10	Q9L111	Q9L111 oryza sativ
37	79.5	6.8	1081	4	Q76065	Q76065 homo sapien
38	79	6.8	220	10	Q9FP95	Q9FP95 oryza sativ
39	79	6.8	224	14	Q9D0D2	Q9D0D2 foot-and-mo
40	79	6.8	224	14	Q9D0D1	Q9D0D1 foot-and-mo
41	79	6.8	268	4	Q9UD7	Q9UD7 homo sapien
42	79	6.8	355	4	Q9UJS5	Q9UJS5 homo sapien
43	79	6.8	365	2	Q52730	Q52730 rhizobium e
44	79	6.8	542	2	Q68872	Q68872 myxococcus
45	79	6.8	866	5	P91426	P91426 caenorhabdi

ALIGNMENTS

RESULT 1
ID Q9Y3U3 PRELIMINARY: PRT: 134 AA.
AC Q9Y3U3:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 15.3 KDA PROTEIN (FRAGMENT).
GN DKFP564D206.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Wambuit R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
EMBL AL050024; CAB43243.1; -.
DR InterPro: IPR000359; -.
DR PROSITE: PS01225; CCKC_2; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 134 AA; 15324 MW; A0B7A8650DZEEGFI CRC64;

Query Match 23.1%; Score 270; DB 4; Length 134;
Best Local Similarity 44.2%; Pred. No. 1.3e-19;
Matches 57; Conservative 23; Mismatches 43; Indels 6; Gaps 4;

QY	80	CRELHFTRYVDGPCRSAPVTELVSGGCGPARLLPNAIGRG---KMW-RPSGDPFRCI 135
DB	3	CRELHFTRYVDGPCRSAPVTELVSGGCGPARLLPNAIGRG---KMW-RPSGDPFRCI 135
QY	136	PRIYRQKRYQLCPGSGEAPRAKRYLVASCKKRLTRFNHOSLAKPFGTEARPPQKRRP 195
DB	63	NKTRFRQLOQDGP-STRYKITVYACKKRYTRQHNESHNESSMPAKPVQHHRE 121
QY	196	RPRA-RSAK 203
DB	122	RRRASKSK 130

```

RESULT 2
Q9PWB0 PRELIMINARY: PRT: 272 AA.
AC Q9PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
OS CER.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Laassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart."
Cur. Biol. 9:931-938(1999).
EMBL: AF139721; AAD51610.1; -
InterPro: IPR000359; -
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT: 1.
SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match
Best Local Similarity 23.1%; Score 106; DB 13; Length 272;
Matches 32; Conservative 20; Mismatches 57; Indels 26; Gaps 5;

QY 45 EPPPELENNK---TNRABNGSRPHHPETKDYSEYSGRELHFRVYDGPGRSAKPV 100
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human CDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024848; BAB15026.1; -
SQ SEQUENCE 168 AA; 19320 MW; D5A4E4E81BDFC0E CRC64;

Db 129 EPPPRKAKKFWDFMLRKNSASEVYLPKTNMHOETCRILPFSQSAHESCEKV-IV 187
QY 101 TELVSCGCGPARLLPNAIGRGKMWPRSGPDR-----CIPDRYARQVQLCPGGEA 153
Db 188 QNNLCFCFKCS-----FHVPGPDRILYTFCSKCLPRTFSMKHLDNCT-SSV 233
QY 154 PRARKVRLVASCCK 168
Db 234 PVKAKMIVECNC 248

RESULT 3
Q9PUK2 PRELIMINARY: PRT: 272 AA.
AC Q9PUK2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CARONTE.
OS CAR.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez Esteban C., Capdevila J., Economides A.N., Pascual J.,
RA Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-1-like protein, mediates the establishment of
left-right asymmetric."
Nat. 400:199-204(1999).
EMBL: AF179484; AAD55581.1; -
InterPro: IPR000359; -
DR InterPro: IPR001839; -
DR Prodom: PD000357; -
DR PROSITE: PS01225; CTCK_2; 1.

```

```

DR SMART: SM00041; CT: 1.
SQ SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match
Best Local Similarity 8.9%; Score 104; DB 13; Length 272;
Matches 32; Conservative 19; Mismatches 58; Indels 26; Gaps 5;

QY 45 EPPPELENNK---TNRABNGSRPHHPETKDYSEYSGRELHFRVYDGPGRSAKPV 100
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human CDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024848; BAB15026.1; -
SQ SEQUENCE 168 AA; 19320 MW; D5A4E4E81BDFC0E CRC64;

Db 129 EPPPRKAKKFWDFMLRKNSASEVYLPKTNMHOETCRILPFSQSAHESCEKV-IV 187
QY 101 TELVSCGCGPARLLPNAIGRGKMWPRSGPDR-----CIPDRYARQVQLCPGGEA 153
Db 188 QNNLCFCFKCS-----FHVPGPDRILYTFCSKCLPRTFSMKHLDNCT-SSV 233
QY 154 PRARKVRLVASCCK 168
Db 234 PVKAKMIVECNC 248

RESULT 4
Q9H772 PRELIMINARY: PRT: 168 AA.
AC Q9H772;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CDNA: FLJ21195 FIS, CLONE COL00185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human CDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024848; BAB15026.1; -
SQ SEQUENCE 168 AA; 19320 MW; D5A4E4E81BDFC0E CRC64;

Query Match
Best Local Similarity 8.8%; Score 103; DB 4; Length 168;
Matches 45; Conservative 26; Mismatches 65; Indels 48; Gaps 9;

QY 2 QULPLALCVCLVHTAFRVVGGQWQAFKNDAREIIPELGEYEPPEPELENNKTNRAEN 61
Db 4 KLSLSEFLVAVLVYA-----EARKNRPA-----GALPSYKDDSSN-----N 41
QY 62 GGRPHHPETKDYSEYSGRELHFRVYDGPGRSAKPVTELVCSCQ 108
Db 42 SERWQHQLKEVLASSQELAVYTERKYLKSDCKTOPLRQTVSEGRS-RTILNFFCYGQ 100
QY 109 CGPARLLPNAIGRGKMWPRSGPDR-----CIPDRYARQVQLCPGGEAP-FARKVRLVA 163
Db 101 CN-SYIPRHV-----KKEEFSCAFCKPQRTSVLVELECPOLDPFFLAKTIQYK 153
QY 164 SCKC 167
Db 154 QCRK 157

RESULT 5
Q9UEM9 PRELIMINARY: PRT: 493 AA.
AC Q9UEM9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:46:18 ; Search time 80.28 seconds
(without alignments)
160.848 Million cell updates/sec

Title: US-09-668-021-2

Perfect score: 1167
Sequence: 1 MQLPLALCVLLVHTAFRV.....KPRPRASANAQAELENAY 213

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0601:*

- 1: /SIDS8/gcgdata/geneseq/AA1980.DAT:*
- 2: /SIDS8/gcgdata/geneseq/AA1981.DAT:*
- 3: /SIDS8/gcgdata/geneseq/AA1982.DAT:*
- 4: /SIDS8/gcgdata/geneseq/AA1983.DAT:*
- 5: /SIDS8/gcgdata/geneseq/AA1984.DAT:*
- 6: /SIDS8/gcgdata/geneseq/AA1985.DAT:*
- 7: /SIDS8/gcgdata/geneseq/AA1986.DAT:*
- 8: /SIDS8/gcgdata/geneseq/AA1987.DAT:*
- 9: /SIDS8/gcgdata/geneseq/AA1988.DAT:*
- 10: /SIDS8/gcgdata/geneseq/AA1989.DAT:*
- 11: /SIDS8/gcgdata/geneseq/AA1990.DAT:*
- 12: /SIDS8/gcgdata/geneseq/AA1991.DAT:*
- 13: /SIDS8/gcgdata/geneseq/AA1992.DAT:*
- 14: /SIDS8/gcgdata/geneseq/AA1993.DAT:*
- 15: /SIDS8/gcgdata/geneseq/AA1994.DAT:*
- 16: /SIDS8/gcgdata/geneseq/AA1995.DAT:*
- 17: /SIDS8/gcgdata/geneseq/AA1996.DAT:*
- 18: /SIDS8/gcgdata/geneseq/AA1997.DAT:*
- 19: /SIDS8/gcgdata/geneseq/AA1998.DAT:*
- 20: /SIDS8/gcgdata/geneseq/AA1999.DAT:*
- 21: /SIDS8/gcgdata/geneseq/AA2000.DAT:*
- 22: /SIDS8/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1167	100.0	213	21	AAB26106 Human DAN/Cerberus
2	1167	100.0	213	21	AAV96429 Human TGF-beta bin
3	1167	100.0	213	22	AAV97589 Human secreted pro
4	1158	99.2	213	21	AAV96436 Human TGF-beta bin
5	1157	99.1	213	21	AAV96430 Human TGF-beta bin
6	1146	98.2	213	21	AAV96431 Vervet TGF-beta bi
7	1072	91.9	367	21	AAB26105 Human DAN/Cerberus
8	1050	90.0	213	21	AAV96433 Rat TGF-beta bindi
9	1029	88.2	211	21	AAV96432 Murine TGF-beta bi
10	901.5	77.2	176	21	AAV96434 Bovine TGF-beta bi
11	350.5	30.0	206	21	AAV75981 Murine skin cell p

12	350.5	30.0	206	21	AAV76031
13	350.5	30.0	206	22	AAV55920
14	350.5	30.0	206	22	AAV55970
15	349.5	29.9	206	18	AAV09408
16	349.5	29.9	206	19	AAV58704
17	349.5	29.9	206	20	AAV95711
18	349.5	29.9	206	21	AAV10233
19	266	22.8	50	20	AAV12009
20	122.5	10.5	116	18	AAV27654
21	122.5	10.5	116	18	AAV44090
22	118	10.1	23	21	AAV96435
23	104.5	9.0	102	20	AAV03225
24	104.5	9.0	102	21	AAV10277
25	103	8.8	168	21	AAV84014
26	98	8.4	158	21	AAV51132
27	96	8.2	184	19	AAV69293
28	96	8.2	184	20	AAV42173
29	96	8.2	184	21	AAV95961
30	90	7.7	184	20	AAV42172
31	90	7.7	184	20	AAV42174
32	90	7.7	429	20	AAV42175
33	89	7.6	184	19	AAV29727
34	88	7.5	558	12	AAV13150
35	87	7.5	391	20	AAV42178
36	86.5	7.4	182	19	AAV29726
37	86.5	7.4	246	21	AAV78333
38	84.5	7.2	561	12	AAV12342
39	82.5	7.1	1480	13	AAV25079
40	82	7.0	150	21	AAV43879
41	82	7.0	267	20	AAV96212
42	82	7.0	267	20	AAV86032
43	82	7.0	270	19	AAV41250
44	82	7.0	562	6	AAV50342
45	82	7.0	562	9	AAV82582

ALIGNMENTS

RESULT 1

ID AAB26106 standard; Protein: 213 AA.

XX AAB26106;

XX 15-JAN-2001 (first entry)

DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).

XX Human: DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;

KW antagonist; BMP; cell growth; cell differentiation; bone formation;

KW gene therapy.

XX Homo sapiens.

OS Homo sapiens.

XX WO200055193-A2.

XX 21-SEP-2000.

XX 02-MAR-2000; 2000WO-US05537.

XX 12-MAR-1999; 99US-0124118.

XX (REG-) REGENERON PHARM INC.

XX Economides AN.

XX WPI: 2000-638179/61.

XX N-PSDB: AAA94051.

XX Novel isolated, human DNA/Cerberus related protein 6 which include

PT natural homologue, and polypeptides comprising DCR6 domain and nucleic

PT acids encoding the proteins which are useful as probes and primers

XX PS Claim 8; Fig 3; 40bp; English.

XX CC The present sequence comprises the amino acid sequence encoded by exons 1

XX CC and 4 of the human DAN/Gerberus-related protein 6 (hDGR6) coding

XX CC sequence. The coding sequence was isolated from a human kidney cDNA

XX CC library containing exons 1 and 4 of the sequence. hDGR6 is closely

XX CC related to the DAN and DGR5 proteins, both of which act as antagonists of

XX CC morphogenic proteins such as BMP. It is possible that the hDGR6 gene and

XX CC protein can be used as immunogens, modulators of cell function, growth

XX CC and differentiation, to reduce undesirable bone formation, to identify

XX CC DGR6 binding agents, in diagnosis, and in gene therapy.

SQ Sequence 213 AA:

Query Match 100.0%; Score 1167; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 3.4e-104;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLPLALCLVCLVHTAFRRVVEGQWQAFKNDATETIIPELGEPPEPPELNNKTMNRAE 60
 1 mqlplalclvcllvhtafrrvveggwqafkndatellpelgeyppelpennktnmrae 60

OY 61 NGRPRPHHPETKDVSEYSCRELHFTRYVTGDFCRSAKPVTELVCSGCGPARLLPNAIG 120
 61 ngrprphhpctkdvsescrelhfttryvtgdfcrsakpvtelvcsgcgparllpnai 120

DB 61 ngrprphhpctkdvsescrelhfttryvtgdfcrsakpvtelvcsgcgparllpnai 120

OY 121 RGMWMPSCGDFRCIPDRYRAQVOLLCPGEGAPRAKRYRLVASCKCKLTFPHNOSELK 180
 121 rgmwmpscgdfrcipdryraqvqlcpgeaprarckryrlvasckckltrfhngselk 180

DB 121 rgmwmpscgdfrcipdryraqvqlcpgeaprarckryrlvasckckltrfhngselk 180

OY 181 DFGTEAARPOKGRKPRPRARSAKANOAELFNAY 213
 181 dfgteaarpqkgrkprprarsakanogelenay 213

DB 181 dfgteaarpqkgrkprprarsakanogelenay 213

RESULT 2
 AAY96429
 ID AAY96429 standard; Protein; 213 AA.

XX AC AAY96429;

XX DT 12-SEP-2000 (first entry)

XX DE Human TGF-beta binding protein (BEER).

XX KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.

XX Homo sapiens.

PN WO200032773-A1.

XX PD 08-JUN-2000.

XX PE 24-NOV-1999; 99WO-US27990.

XX PR 27-NOV-1998; 98US-0110283.

XX PA (DARWIN) DARWIN DISCOVERY LTD.

PI Brunhlow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;

XX DR WPI; 2000-412321/35.

XX DR N-PSDB; AAA29055.

XX PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

XX Claim 2; Page 116; 162pp; English.

XX CC This shows the human transforming growth factor-beta (TGF-beta)

XX CC binding protein designated hBEER. The cDNA and protein may be used for

XX CC prevention, treatment and diagnosis of diseases associated with

XX CC inappropriate BEER expression. For example, they may be used to treat

XX CC disorders associated with decreased TGF-beta BP expression. The cDNA or

XX CC vectors may be administered to treat diseases by rectifying mutations or

XX CC deletions in a patient's genome that affect the activity of BEER by

XX CC expressing inactive proteins or to supplement the patient's own production

XX CC of BEER polypeptides. The nucleic acids may be used for recombinant

XX CC production of BEER, gene therapy, antisense therapy, as probes for

XX CC diagnostic assays and for functional studies. BEER may be used to raise

XX CC antibodies and for identification of BEER modulators. BEER antagonists

XX CC may be used to increase bone mineral content for the treatment of

XX CC disorders such as osteopenia, osteoporosis, fractures and other

XX CC disorders associated with low mineral content.

SQ Sequence 213 AA:

Query Match 100.0%; Score 1167; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 3.4e-104;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MQLPLALCLVCLVHTAFRRVVEGQWQAFKNDATETIIPELGEPPEPPELNNKTMNRAE 60
 1 mqlplalclvcllvhtafrrvveggwqafkndatellpelgeyppelpennktnmrae 60

DB 1 mqlplalclvcllvhtafrrvveggwqafkndatellpelgeyppelpennktnmrae 60

OY 61 NGRPRPHHPETKDVSEYSCRELHFTRYVTGDFCRSAKPVTELVCSGCGPARLLPNAIG 120
 61 ngrprphhpctkdvsescrelhfttryvtgdfcrsakpvtelvcsgcgparllpnai 120

DB 61 ngrprphhpctkdvsescrelhfttryvtgdfcrsakpvtelvcsgcgparllpnai 120

OY 121 RGMWMPSCGDFRCIPDRYRAQVOLLCPGEGAPRAKRYRLVASCKCKLTFPHNOSELK 180
 121 rgmwmpscgdfrcipdryraqvqlcpgeaprarckryrlvasckckltrfhngselk 180

DB 121 rgmwmpscgdfrcipdryraqvqlcpgeaprarckryrlvasckckltrfhngselk 180

OY 181 DFGTEAARPOKGRKPRPRARSAKANOAELFNAY 213
 181 dfgteaarpqkgrkprprarsakanogelenay 213

DB 181 dfgteaarpqkgrkprprarsakanogelenay 213

RESULT 3
 AAY97589
 ID AAY97589 standard; Protein; 213 AA.

XX AC AAY97589;

XX DT 05-APR-2001 (first entry)

XX DE Human secreted protein PRO7476.

XX KM secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KM proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KM angiogenic disorder; immunologic disorder; PRO7476.

XX Homo sapiens.

PN WO200075317-A2.

XX PD 14-DEC-2000.

XX PE 15-MAY-2000; 2000WO-US13358.

XX PR 09-JUN-1999; 99US-0138385.

XX PR 20-JUL-1999; 99US-0144790.

XX PR 03-AUG-1999; 99US-0146843.

XX PR 10-AUG-1999; 99US-0148188.

XX PR 17-AUG-1999; 99US-0149320.

XX PR 17-AUG-1999; 99US-0149327.

XX PR 17-AUG-1999; 99US-0149396.

XX PR 20-AUG-1999; 99US-0150114.

XX PR 31-AUG-1999; 99US-0151700.

XX PR 31-AUG-1999; 99US-0151734.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:49:43 ; Search time 54.26 Seconds

(without alignments)
299.027 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 1165

Sequence: 1 MQLPLALCLICLLVHTAFRV.....KPPRRARSAAKQALEENAY 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR-68:*

1: pirl:*

2: pirl2:*

3: pirl3:*

4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270	23.2	134	2 T08710	hypothetical prote
2	94.5	8.1	159	2 I51373	luteinizing hormon
3	87.5	7.5	158	2 A61091	lutropin beta chai
4	87.5	7.5	601	2 T22025	hypothetical prote
5	84.5	7.3	824	2 T10020	leucine--trna 11ga
6	84.5	7.3	972	2 T10023	leucine--trna 11ga
7	84	7.2	1469	2 B36665	slit protein 2 pre
8	82	7.0	270	2 S71793	head-inducing fact
9	81.5	7.0	1480	2 A36665	slit protein 1 pre
10	80.5	6.9	969	2 A70912	probable lens prot
11	80	6.9	510	2 A42750	insulinoma-associ
12	80	6.9	839	2 I50590	class I INCEMP pro
13	80	6.9	866	2 T29197	hypothetical prote
14	80	6.9	2531	2 T16743	hypothetical prote
15	79.5	6.8	1122	2 T14180	exit protein - Myc
16	79	6.8	118	2 PN0141	lutropin beta chai
17	79	6.8	178	2 I51824	tumor-suppressive
18	79	6.8	178	2 A47291	probable finger pr
19	79	6.8	243	2 S28444	lomb protein - pse
20	79	6.8	2333	1 GNNY2F	genome polyprotein
21	78.5	6.7	166	2 I51242	luteinizing hormon
22	78.5	6.7	1138	2 G85077	probable polyprote
23	78	6.7	646	2 T02398	hypothetical prote
24	78	6.7	1142	2 T30272	hypothetical prote
25	78	6.7	1210	1 GQHU6	epidermal growth f
26	77.5	6.7	343	2 G35070	apolipoprotein H-r
27	77.5	6.7	864	2 H85335	hypothetical prote
28	77.5	6.7	864	2 T04518	hypothetical prote
29	77.5	6.7	1042	2 A57534	mucln 5AC (clone L

30	77.5	6.7	2142	2 B35098	MHC class III hist
31	77	6.6	118	2 PN0139	lutropin beta chai
32	77	6.6	520	2 S47142	matrig type A prot
33	77	6.6	613	2 S15468	complement C3b/C4b
34	77	6.6	702	2 E72775	probable helicase
35	77	6.6	877	2 I50591	class II INCEMP pr
36	77	6.6	992	2 A31666	hypothetical prote
37	76.5	6.6	402	2 T48729	hypothetical prote
38	76.5	6.6	715	2 T12534	hypothetical prote
39	76.5	6.6	1223	2 S29717	adenylate cyclase
40	76	6.5	398	2 A35281	integratory muci
41	76	6.5	562	1 UKH0T	t-plasminogen acti
42	75.5	6.5	403	2 C70832	hypothetical prote
43	75.5	6.5	543	1 JC4070	protein kinase (EC
44	75.5	6.5	1436	2 A46496	antigen WC1.1 prec
45	74	6.4	613	2 T00077	gag-like protein -

ALIGNMENTS

RESULT 1

T08710 hypothetical protein DKFP564D206.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence-revision 11-Jun-1999 #text-change 13-Aug-1999

C:Accession: T08710

R:Wambuit, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999

A:Reference number: 216471

A:Accession: T08710

A:Molecule type: mRNA

A:Residues: 1134 <NAM>

A:Cross-references: EMBL:AL050024

A:Experimental source: Fetal brain; clone DKFP564D206

C:Genetics:

A>Note: DKFP564D206.1

Query Match 23.2%; Score 270; DB 2; Length 134;

Best Local Similarity 44.2%; Pred. No. 3.4e-17;

Matches 57; Conservative 23; Mismatches 43; Indels 6; Gaps 4;

QY 80 CRELHPTRYVTDGCRARKEPVTEVLCGCGGPARLLPNAIGRG---KMW-RESPDFPCI 135

DB 3 CRELRSTKIYSDQCTISPLKELVAGCELLPVPINMGSGYGTIKYSRKSQEMKCV 62

QY 136 PDYRAORVOLLCPGEPAPARKVRLVASCKKRLTRFHNOSLDFGTEARPOKGRKP 195

DB 63 NDKTRTRIOLOQODG-STRTYKITVTACKCKRYRQHNESHNEFSWSPAKPVQHRE 121

QY 196 RPRR-RSAK 203

DB 122 RKRASKSK 130

RESULT 2

I51373 luteinizing hormone beta subunit - turkey

C:Species: Melaleuca gallipavo (common turkey)

C>Date: 13-Sep-1996 #sequence-revision 13-Sep-1996 #text-change 16-Jul-1999

C:Accession: I51373

R:You, S.; Foster, L.K.; Silsby, J.D.; el Halawani, M.E.; Foster, D.N.

J. Mol. Endocrinol. 14, 117-129, 1995

A:Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonado

A:Reference number: I51373; MUID:95290073

A:Accession: I51373

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-159 <Y0>

A:Cross-references: GB:IJ5519; NID:9530952; PIDN:AAA74125.1; PID:9530953

C:Genetics:

A:Gene: LH-beta

C;Superfamily: pituitary glycoprotein hormone beta chain

Query Match	8.1%;	Score 94.5;	DB 2;	Length 159;
Best Local Similarity	29.4%;	Pred. No. 0.22;		
Matches 32;	Conservative 7;	Mismatches 49;	Indels 21;	Gaps 5;

```
Qy      62 GGRPHHPEETKDVSEYSCRELHFTRYVTDGPCHSAKPVTELVCSOGCGPARLLPNAIGR 121
        |||| |::| | | | | : |
Db      43 GGRP-----CRPIWTVAAVEKDECFQMAYTTACGTCRT---EPYR 85
```

Qy 122 GKWWRPSGPDFRCIPDRYRAQRVQL-CPGEAPRARKVRVLVASCCKR 169
||| :||| :|||
Db 86 SPLGRP-PQSSCTYGALRYERWALMGCPISDPRV-LLPVALSCRCAR 131

RESULT 3
A61091
lutropin beta chain precursor - chicken
N;Alternate names: luteinizing hormone beta chain

Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995
 Accession: A61091
 Molecule: T.; Ando, H.; Ueda, T.; Kubokawa, K.; Higashinakagawa, T.; Ishii, S.
 J. Mol. Endocrinol. 3, 129-137, 1989
 Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for the
 Reference number: A61091; MUID:89374710
 Accession: A61091
 A:Status: not compared with conceptual translation
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-158 <NOC>
 C:Superfamily: pituitary glycoprotein hormone beta chain
 C:Keywords: glycoprotein; hormone; pituitary
 F:1-39/domain: signal sequence #status predicted <SIG>
 F:48-73, 62-96, 65-127, 77-149, 111-139, 129-132/disulfide bonds: #status predicted
 F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match      7.5%; Score 87.5; DB 2; Length 158;
Best Local Similarity 27.9%; Ped. No. 0.93;
Matches 31; Conservative 8; Mismatches 47; Indels 25; Gaps 5.

Oy 62 GGRPPHPPETKDVSEYSCRELFHFRYYTGDGCRSARPYTELVCSGGCGPARLLPNAIGH 121
    |||||
Db 43 GGRPP-----CRPIINVAAEKDDCCPCMAVTTTACGGGCKRTREVP----- 83

```

```

QY      122 GKWWRPSC--PDFRCIPDRYRAQRVQLL-CPGGEAPRARARYLVASCKCKR 169
      : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      84 --YRSPFLGPPQSACTYGAALRYERWALWGCPIGSDPRV-LLEVALSCRCAR 131

```

SUIT 4
 T22025
 Hypothetical protein F40E10.4 - *Caenorhabditis elegans* (fragment)
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T22025
 R:Smyle, R.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19503
 A:Accession: T22025
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-601 <NTL>
 A:Cross-references: EMBL:Z69792; PIDD:CAA9366.1; GSPDB:GND00028; CESP:F40E10.4
 A:Experimental source: Clone F40E10
 C:Genetics:
 A:Gene: CESP:F40E10.4
 A:Map position: X

Query Match	7.58;	Score 87.5;	DB 2;	Length 601;
Best Local Similarity	20.68;	Pred. No. 3.6;		

Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

67 HHPETKDVSEYSCRELHFTRYIVTDGPPCRBAKPVTELVCSGGCGPARLLPNAIGRGKWNR 126
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
510 HFSGEHCDKERIKCDKQKFRHHIENECRSRDRIKIACNGNYCG----- 553

127 PGGPDERCIDRYRAORVOLLCPGGEAPRARKVRLVASCKCK 168
| : :::: | : | :: 1:1:
554 --GEQNCSTAVKKKQKQVKMKCKNG--TTKTSVTHIRQQCE 592

RESULT 5
110020
eucine--tRNA Ligase (EC 6.1.1.4) - Mycobacterium leprae (fragment)

```

:::Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
:::Accession: T10020
:::Cole, S.T.
:::Submitted to the EMBL Data Library, August 1997
:::Reference number: Z16916
:::Accession: T10020
:::Status: preliminary; translated from GB/EMBL/DBJ
:::Molecule type: DNA
:::Residues: 1-824 <COL>
:::Cross-references: EMBL:Z70722
:::Genetics:
:::Gene: leuS,
:::Superfamily: leucine--tRNA ligase
:::Keywords: aminacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match      7.3%; Score 84.5; DB 2; Length 824;
Best Local Similarity 26.0%; Pred. No. 9.1;
Matches 33; Conservative 15; Mismatches 76; Indels 17; Gaps 6;

```

```

16 TAFRVVEGGQWQAKNDATETIRELGEXPEBPPELNNKTMNRAENG----GRPHHPEE 7
      | : | | : : | : | : | : |
106 TGHNVLHAMGFDAEGFLPAEQYAMQTGTHRILTEA--NVNFRQLGLGLGHSRRFTS 1

```

72 TKDSEYSCRELHFTTRYVT--DGPCRSAPVTELYCSGCCPARLLPNAIGRCKWRPS 128
 164 TTDFVEFYKWTQWIFLQIYNAMEDVANKARPIAELIAEFDSGERRLYD--GR-DNATLLS 219

```

129 GPDFRCIPDR---YRAQRVQLCPG 150
      : : | | | : | | |
220 AGERADVIDNCRLVYRADSMVNWCPG 245

```

```

RESULT 6
10023
leucine--tRNA leucine--tRNA ligase (EC 6.1.1.4) - Mycobacterium leprae
Species: Mycobacterium leprae
Accession: T10023
Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
Author: Eiglmeier, K.; Honore, N.; Woods, S.A.; Caudron, B.; Cole, S.T.
J Mol Microbiol. 7, 197-206, 1993
Title: Use of an ordered cosmid library to deduce the genomic organization of Mycob
Reference number: Z16917, MIMD:93188700
Accession: T10023
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 11972 <EIG>
Cross-references: EMBL:Y14967; NID:g2370268; PIDD:CAAF5192.1; PID:g2370271
Genetics:
Gene: leuS
Superfamily: leucine--tRNA ligase
Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

```

Query Match	7.38;	Score 84.5;	DB 2;	Length 972;
Best Local Similarity	26.0%;	Pred. No. 11;		
Matches	38;	Conservative	15;	Mismatches 76;
				Indels 17;
				Gaps 6.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:12 ; Search time 31.13 seconds
(without alignments)
234.385 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 1165
Sequence: 1 MQLPLALCLICLLVHTAFRV.....KPPRRASAKANAELEENAY 213

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

arched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94.5	8.1	159	1	LSHB_MELGA
2	85	7.3	180	1	DAN_HUMAN
3	84.5	7.3	972	1	SYL_MYCLE
4	83.5	7.2	141	1	LSHB_TIRIV
5	81.5	7.0	1480	1	SLIT_DROME
6	80.5	6.9	969	1	SYL_MYCTU
7	80	6.9	530	1	IAT1_HUMAN
8	79.5	6.8	318	1	HK25_MOUSE
9	79	6.8	118	1	LSHB_PRYCA
10	79	6.8	178	1	DAN_RAT
11	79	6.8	243	1	TOMB_PSEPU
12	79	6.8	355	1	KLP2_HUMAN
13	79	6.8	870	1	BCA1_HUMAN
14	79	6.8	2333	1	POLG_FMDVI
15	78.5	6.7	166	1	LSHB_COTVA
16	78	6.7	958	1	HIG_DROME
17	78	6.7	1210	1	EGFR_HUMAN
18	77.5	6.7	951	1	SFR8_HUMAN
19	77.5	6.7	2142	1	BAT2_HUMAN
20	77	6.6	118	1	LSHB_BALAC
21	77	6.6	877	1	INCE_CHICK
22	77	6.6	992	1	EBN6_EBV
23	76.5	6.6	354	1	ATH1_HUMAN
24	76	6.5	112	1	PSC2_RAT
25	76	6.5	398	1	MU81_XENLA
26	76	6.5	562	1	TPA_HUMAN
27	75.5	6.5	318	1	HK25_RAT
28	75.5	6.5	543	1	PRAA_SPROCO
29	75.5	6.5	1436	1	WC11_BOVIN
30	75	6.4	141	1	LSHB_CERST
31	74	6.4	138	1	LSHB_MACRU
32	74	6.4	470	1	KLP4_HUMAN
33	74	6.4	544	1	H15_DROME

34	74	6.4	2715	1	TRX2_HUMAN
35	73.5	6.3	301	1	UL49_HSV11
36	73.5	6.3	682	1	SNK_MOUSE
37	73.5	6.3	682	1	SNK_RAT
38	73.5	6.3	685	1	SNK_HUMAN
39	73	6.3	178	1	DAN_MOUSE
40	73	6.3	191	1	CBX5_HUMAN
41	73	6.3	972	1	POLS_IPNVN
42	72.5	6.2	128	1	LSHB_STRCA
43	72.5	6.2	375	1	CE10_CHICK
44	72.5	6.2	755	1	RRE1_HUMAN
45	72	6.2	379	1	CYR6_MOUSE

ALIGNMENTS

RESULT 1	
LSHB_MELGA	STANDARD; PRT; 159 AA.
AC P45646;	
DT 01-NOV-1995 (Rel. 32, Created)	
DT 01-NOV-1995 (Rel. 32, Last sequence update)	
DT 01-OCT-2000 (Rel. 40, Last annotation update)	
DE LUTROPIN BETA CHAIN PRECURSOR (LUTHEINIZING HORMONE BETA SUBUNIT) (LSH- BETA) (LSH-B) (LH-B).	
GN LHB.	
OS Meleagris gallopavo (Common turkey).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.	
OX NCBI_TaxID=9103;	
RN [1]	
RE SEQUENCE FROM N.A.	
RC TISSUE=Plutitary;	
RX MEDLINE=95290073; PubMed=7772235;	
RA You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;	
RT "Sequence analysis of the turkey LH beta subunit and its regulation by gonadotrophin-releasing hormone and prolactin in cultured pituitary cells.";	
RT J. Mol. Endocrinol. 14:117-129(1995).	
RU -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.	
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.	
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).	
CC	
CC EMBL; L35519; AAA74125.1; ALT_INIT.	
DR HSSP; P01233; IHRP.	
DR InterPro; IPR000359; -	
DR InterPro; IPR001545; -	
DR InterPro; IPR002400; -	
DR Pfam; PF00007; Cys_knot; 1.	
DR PRINTS; PR00438; GFCYSKNOT.	
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.	
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.	
FW Hormone; Signal; Glycoprotein.	
FT SIGNAL 1 39	
FT CHAIN 40 159	
FT LUTROPIN BETA CHAIN.	
FT DISULFID 48 96	
FT BY SIMILARITY.	
FT DISULFID 62 111	
FT BY SIMILARITY.	
FT DISULFID 65 149	
FT BY SIMILARITY.	
FT DISULFID 73 127	
FT BY SIMILARITY.	
FT DISULFID 77 129	
FT BY SIMILARITY.	

FT	DISULFID	132	139	BY SIMILARITY.
FT	CARBOHYD	52	52	N-LINKED (GLCNAC...) (POTENTIAL).
SQ	SEQUENCE	159 AA;	16285 MW;	5250C879653C6 CRG64;
Query Match				
	Best Local Similarity	29.4%;	Pred. No.0.051;	Length 159;
	Matches	32;	Conservative	7; Mismatches 49; Indels 21; Gaps 5.
QY	62	GGRRPHHPETKDVSEYSCRELHFRVYLDGFCRSKAPVTELVGSGCGSARLLPNAGR	121	
			:::	:::
Db	43	GGRRP-----CRPIWTVVAVEKDECCPOCMATVTTACGGTCRTR---EPVTR	85	
QY	122	GKMWRFSGPDERCIPDRYRAQRVOLL-CPGSGEAPRARRKRVLASCCKKR	169	
			:::	:::
Db	86	SPICRP--RPSCTTYGALRYERNALMGCPISGSDPVV-LILPALSCRRAR	131	

RESULT 2			
ID	DAN_HUMAN	STANDARD:	PRT: 180 AA.
CC	P41271;		
CC	01-FEB-1995 (Rel. 31, Created)		
DT	01-FEB-1995 (Rel. 31, Last sequence update)		
DT	01-OCT-2000 (Rel. 40, Last annotation update)		
DE	NEUROBLASTOMA SUPPRESSOR OF TUMORIGENICITY 1 (ZINC FINGER PROTEIN DAN) (N03).		
CC	NBL1 OR DAN.		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxId:9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RT	TISSUE=lung;		
RX	MEDLINE-94366724; PubMed-8084583;		
RA	Enomoto H., Ozaki T., Takahashi E., Nomura N., Tabata S.,		
RA	Takahashi H., Ohnuma N., Tanabe M., Iwai J., Yoshida H., Matsunaga T.,		
RA	Sakiyama S.;		
RT	"Identification of human DAN gene, mapping to the putative		
RT	neuroblastoma tumor suppressor locus.";		
RL	Oncogene 9:2785-2791(1994).		
CC	-1- FUNCTION: POSSIBLE CANDIDATE AS A TUMOR SUPPRESSOR GENE OF		
CC	NEUROBLASTOMA. MAY PLAY AN IMPORTANT ROLE IN PREVENTING CELLS		
CC	FROM ENTERING THE FINAL STAGE (G1/S) OF THE TRANSFORMATION		
CC	PROCESS.		
CC	-1- TISSUE SPECIFICITY: MOST ABUNDANT IN NORMAL LUNG AND MENINGIOMA.		
CC	-1- DISEASE: THE LOSS OF DAN GENE BY MUTATION IS POSSIBLY THE CAUSE OF		
CC	THE DEVELOPMENT AND/OR PROGRESSION OF HUMAN NEUROBLASTOMA.		
CC	-1- SIMILARITY: HIGH, TO OTHER MAMMALIAN DAN.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	at the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; D28124; BAA05671.1; -.		
DR	MIM; 600613; -.		
KW	Anti-oncogene; DNA-binding; Zinc-finger.		
FT	ZN FING 81 101 C4-TYPE (POTENTIAL).		
FT	DOMAIN 141 171 PRO-RICH.		
SO	SEQUENCE 180 AA; 19277 MW; 1523352290B655865 CRC64;		

```
Query Match:          7.3%; Score 85; DB 1; Length 180;
Best Local Similarity 26.0%; Pred. No. 0.45;
Matches    33; Conservative   15; Mismatches   61; Indels   18; Gaps    7.
```

Db 34 CEAKNTQIVGHSGC-EAKSIQNRACLGQGC-SYSVNPF-----PQSTSLVHNSC 84
 QY 135 IPDPRYARQRIQLCC-GEAPRARK-VRLVNSACSKRLRFHNSQLKPF--GTEAPRQ 190
 Db 85 MPNSMMELVTLECPRIEVRVQDLVYKTLINCSQACSGKPRSHESLSVYVGEGGPGSQ 144
 QY 191 KGRKPRP 197
 Db 145 PGTINPF 151

RESULT	3
SYL_MYCLE	
ID	SYL_MYCLE
AC	Q50192;
DT	15-JUL-1998 (Rel. 36, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	LEUCYL--TRNA SYNTHETASE (EC 6.1.1.4) (LEUCINE--TRNA LIGASE) (LEURS).
GN	LEUC.
OS	Mycobacterium leprae.
OC	Bacteria, Firmicutes, Actinobacteria, Actinobacteridae;
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX	NB01_taxid=1769;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Egglemeier K., Garnier T., de Rossi E., Fajni H., Cole S.T.;
RL	Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC	-1- CATALYTIC ACTIVITY: ATP + L-LEUCINE + TRNA(LEU) = AMP +
CC	PYROPHOSPHATE + L-LEUCYL--TRNA(LEU).
CC	-1- SUBCELLULAR LOCATION: CYTOSOL;PLASMIC.
CC	-1- SIMILARITY: BELONGS TO CLASS=1 AMINOACYL--TRNA SYNTHETASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; Y14967; CAA75192.1; -;
DR	EMBL; Z70722; CAA94727.1; -;
DR	InterPro; IPR001412; -;
DR	InterPro; IPR002300; -;
DR	InterPro; IPR002302; -;
DR	Pfam; PF00133; tRNA-synt_1_1;
DR	PRINTS; PR00985; TRNASYNTHL1E.
DR	PROSITE; PS00178; AA-TRNA_LIGASE_1; 1.
KW	Aminoacyl--trna synthetase; Protein biosynthesis; Ligation; ATP-binding.
FT	1
FT	2
FT	3
FT	4
FT	5
FT	6
FT	7
FT	8
FT	9
FT	10
FT	11
FT	12
FT	13
FT	14
FT	15
FT	16
FT	17
FT	18
FT	19
FT	20
FT	21
FT	22
FT	23
FT	24
FT	25
FT	26
FT	27
FT	28
FT	29
FT	30
FT	31
FT	32
FT	33
FT	34
FT	35
FT	36
FT	37
FT	38
FT	39
FT	40
FT	41
FT	42
FT	43
FT	44
FT	45
FT	46
FT	47
FT	48
FT	49
FT	50
FT	51
FT	52
FT	53
FT	54
FT	55
FT	56
FT	57
FT	58
FT	59
FT	60
FT	61
FT	62
FT	63
FT	64
FT	65
FT	66
FT	67
FT	68
FT	69
FT	70
FT	71
FT	72
FT	73
FT	74
FT	75
FT	76
FT	77
FT	78
FT	79
FT	80
FT	81
FT	82
FT	83
FT	84
FT	85
FT	86
FT	87
FT	88
FT	89
FT	90
FT	91
FT	92
FT	93
FT	94
FT	95
FT	96
FT	97
FT	98
FT	99
FT	100
FT	101
FT	102
FT	103
FT	104
FT	105
FT	106
FT	107
FT	108
FT	109
FT	110
FT	111
FT	112
FT	113
FT	114
FT	115
FT	116
FT	117
FT	118
FT	119
FT	120
FT	121
FT	122
FT	123
FT	124
FT	125
FT	126
FT	127
FT	128
FT	129
FT	130
FT	131
FT	132
FT	133
FT	134
FT	135
FT	136
FT	137
FT	138
FT	139
FT	140
FT	141
FT	142
FT	143
FT	144
FT	145
FT	146
FT	147

	Query Match	7.3%	Score 84.5;	DB 1;	Length 972;	
	Best Local Similarity	26.0%	Pred. No. 3;			
	Matches	38;	Conservative	15;	Mismatches	76;
					Indels	17;
					Gaps	6;
QY	16	TAFRVVGGGMAFKNDATETIRELGEYFEPPELENNKTYNRAENG---	GRPHHFE	71		
Db	106	TGNHVAMGEDAFLGAEDYAMQGTHIRILTEA--NVVNFRHQRLSLGIHDSRTFS	163			
QY	72	TKDYSEISCELTHTFTRYT---DGPCKSKPYTELVCSGCGCPARLLPNAIGRKKWRRPS	128			
Db	164	TTDFEFKWTQMFILOIYNAMEDVANAKRKAPITAEIIAEFDGERRIVD---GR-DWATLS	219			
QY	129	GDPFCIPDR---YRAQRVOLLCEG	150			
Db	220	AGERADYIDNCRLVYRADSKVMWCFFS	245			

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:00:34 ; Search time 89.34 Seconds

(without alignments)
315.435 Million cell updates/sec

Title: US-09-668-021-6

Sequence: 1 MOLPLALICLLVHTAFRV.....KPRPRASAKAQAELNNY 213

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 425026 segs, 13305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPTREMBL_16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270	23.2	134	4 Q9Y3U3	Q9Y3U3 homo sapien
2	108	9.3	168	4 Q9H772	Q9H772 homo sapien
3	106	9.1	272	13 Q9PWB0	Q9PWB0 gallus gall
4	104	8.9	272	13 Q9PWB0	Q9PWB0 gallus gall
5	103	8.8	168	11 Q9H8273	Q9H8273 mus musculu
6	98.5	8.5	493	4 Q9UEM9	Q9UEM9 homo sapien
7	94.5	8.1	493	4 Q9H8273	Q9H8273 mus musculu
8	93.5	8.0	493	4 Q9H8273	Q9H8273 mus musculu
9	93	8.0	1114	11 Q9H8273	Q9H8273 mus musculu
10	92	7.9	1114	11 Q9H8273	Q9H8273 mus musculu
11	91	7.8	215	8 Q9MFB6	Q9MFB6 beta vulgar
12	90	7.7	184	11 Q9H8273	Q9H8273 mus musculu
13	88.5	7.6	492	11 Q9H8273	Q9H8273 mus musculu
14	88.5	7.5	184	13 Q9H8273	Q9H8273 mus musculu
15	87.5	7.4	182	13 Q9H8273	Q9H8273 mus musculu
16	86.5	7.1	661	5 Q9V708	Q9V708 xenopus lae
17	82.5	7.1	791	11 Q9H8273	Q9H8273 mus musculu
18	82.5	7.0	270	13 P70041	P70041 xenopus lae
19					

20	82	7.0	988	6 Q97867	Q97867 sus scrofa
21	81.5	7.0	1480	5 Q9V7E8	Q9V7E8 drosophila
22	81.5	7.0	1504	5 Q9XYV4	Q9XYV4 drosophila
23	81.5	7.0	1504	5 Q9V7E9	Q9V7E9 drosophila
24	81.5	7.0	355	4 Q9UKR6	Q9UKR6 homo sapien
25	80.5	6.9	267	4 Q9S813	Q9S813 homo sapien
26	80	6.9	712	4 Q9Y557	Q9Y557 homo sapien
27	80	6.9	720	4 Q9N1Z3	Q9N1Z3 homo sapien
28	80	6.9	866	5 P91426	P91426 caenorhabdi
29	80	6.9	1440	5 Q92024	Q92024 caenorhabdi
30	80	6.9	2531	5 Q21980	Q21980 caenorhabdi
31	79.5	6.8	500	10 Q9M6C2	Q9M6C2 trifolium r
32	79.5	6.8	1081	4 Q76065	Q76065 mus sapien
33	79.5	6.8	1122	2 Q85018	Q85018 mycobacteri
34	79.5	6.8	1181	5 Q9V795	Q9V795 drosophila
35	79	6.8	220	10 Q9P995	Q9P995 oryza sativ
36	79	6.8	329	2 Q9W360	Q9W360 acetobacter
37	79	6.8	355	4 Q9UJ55	Q9UJ55 homo sapien
38	78.5	6.8	542	2 Q68872	Q68872 myxococcus
39	78.5	6.7	240	10 Q9SNN5	Q9SNN5 oryza sativ
40	78.5	6.7	365	2 Q52730	Q52730 rhizobium e
41	78.5	6.7	482	14 Q71654	Q71654 human immun
42	78.5	6.7	1096	10 Q9XE24	Q9XE24 oryza sativ
43	78.5	6.7	1138	10 Q9ZPF9	Q9ZPF9 arabidopsis
44	78	6.7	224	14 Q9DUD2	Q9DUD2 foot-and-mo
45	78	6.7	224	14 Q9DUD1	Q9DUD1 foot-and-mo

ALIGNMENTS

RESULT 1
ID Q9Y3U3 PRELIMINARY: PRT: 134 AA.
AC -Q9Y3U3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 15.3 KDA PROTEIN (FRAGMENT).
GN DKFZP564D206.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Wamboldt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY 1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL050024; CAB43243.1; .
DR InterPro: IPR000359; .
DR PROSITE: PS01225; CTCK_2; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 134 AA: 15324 MW: A0B7A8650D2EE6F1 CRC64;

Query Match 23.2% Score 270; DB 4; Length 134;
Best local similarity 44.2% Pred. No 1.7e-19;
Matches 57; Conservative 23; Mismatches 43; Indels 6; Gaps 4;

QY 80 CRELHTRVVDGPGRSAPVTELVCSGCGPARLLPNAIGRG---KMW-RPSGDFRCI 135
DB 3 CRELSTRKYSIDGQCTSIISPLKEVCAEGCILLFVLPNWTGGYGRKYSRKSSEWIKV 62
QY 136 PDYRQVROQLCPGGEPRARAKVLAASCKRLTRFNHQSLEKDFGEARLPQGRKP 195
DB 63 NDKTRIRQLQCGQCG-STRTRYKITVYTAACKCKRTPHSHNEHSNPAKPVQHIRE 121
QY 196 RPRRA-RSAK 203
DB 122 RRRASKSK 130

```

RESULT 2
Q9H772 PRELIMINARY; PRT; 168 AA.
ID Q9H772;
AC Q9H772;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CUNA: FLJ21195 FIS, CLONE COL00185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "MEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
EMBL: AF024848; BAB15026.1;
SEQUENCE 168 AA; 19320 MW; D5A4E4E818BF8C0E CRC64;

Query Match 9.3%; Score 108; DB 4; Length 168;
Best Local Similarity 24.5%; Pred. No. 0.0032;
Matches 45; Conservative 27; Mismatches 64; Indels 48; Gaps 9;

QY 2 QLPALCLICLVHTAFRVVGGQGAFAKNDATETIRELCEYPPPELENNKTMNRAEN 61
DB 4 KLSLFLVAVLVKVA-----EARKN-----RPAGAIIPSPYKDGSSN-----N 41

QY 62 GGRPRHPFETKQVSYK-----CRELHFRVYTDGPRCSAKPVLEIVCSQ 108
DB 42 SEMQKQKEFVLASSQSEALVTERKYLKSDCKTQPLQVYSEGCSS-RITLNRFCYQ 100

QY 109 CGPARLLPNAIGRGKWRPSPDFR-----CIPDRYRAQVOLLCPGGEAP-RAKRYVLA 163
DB 101 CN-SFYIPRHV-----KKEEESFQSCAFCKFORVTSVLVELCEPGDLPFFRLKIKQKVA 153

QY 164 SKCK 167
DB 154 QCRC 157

RESULT 3
Q9PWB0 PRELIMINARY; PRT; 272 AA.
ID Q9PWB0;
AC Q9PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart."
RL Curr. Biol. 9:931-938(1999).
DR EMBL: AF139721; AAD51610.1;
DR InterPro: IPR000359;
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT; 1; 1224 MW; 36E4C9F719711BCA CRC64;
SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

```

```

Query Match 9.1%; Score 106; DB 13; Length 272;
Best Local Similarity 23.7%; Pred. No. 0.0084;
Matches 32; Conservative 19; Mismatches 57; Indels 26; Gaps 5;

QY 45 EPPPELENNK-----TNNRANGRRPHHPFETKDYSEKSELHFTRYVTDGCRSAKRY 100
DB 129 EPPRKDAKKFMDHFLRNNSAEEVLLPIKTNEHDETCTLPFSQSVAHSECEKV-IV 187

QY 101 TELVCSGQCGPARLLPNAIGRGKWRPSPDFR-----CIPDRYRAQVOLLCPGGEA 153
DB 188 ONNLCEFGKSS-----FHVGPDDRLYTFCKSKLPYFSMKHLDNCT-SSV 233

QY 154 PRAKRYLVASCKCK 168
DB 234 PVYKKWIVVECNCE 248

RESULT 4
Q9PUK2 PRELIMINARY; PRT; 272 AA.
ID Q9PUK2;
AC Q9PUK2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CARONTE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez Esteban C., Capdevilla J., Economides A.N., Pascual J.,
RA Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-lyk-like protein, mediates the establishment of
embryonic left-right asymmetry."
RL Nature 0:0-0(1999).
DR EMBL: AF179484; AAD55581.1;
DR InterPro: IPR000359;
DR InterPro: IPR001839;
DR PRODOM: PD000357;
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT; 1;
SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match 8.9%; Score 104; DB 13; Length 272;
Best Local Similarity 23.7%; Pred. No. 0.013;
Matches 32; Conservative 19; Mismatches 58; Indels 26; Gaps 5;

QY 45 EPPPELENNK-----TNNRANGRRPHHPFETKDYSEKSELHFTRYVTDGCRSAKRY 100
DB 129 EPPRKDAKKFMDHFLRNNSAEEVLLPIKTNEHDETCTLPFSQSVAHSECEKV-IV 187

QY 101 TELVCSGQCGPARLLPNAIGRGKWRPSPDFR-----CIPDRYRAQVOLLCPGGEA 153
DB 188 ONNLCEFGKSS-----FHVGPDDRLYTFCKSKLPYFSMKHLDNCT-SSV 233

QY 154 PRAKRYLVASCKCK 168
DB 234 PVYKKWIVVECNCE 248

RESULT 5
Q88273 PRELIMINARY; PRT; 168 AA.
ID Q88273;
AC Q88273;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:41 ; Search time 80.28 Seconds
(without alignments)
160.848 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 1165
Sequence: 1 MQLPLALCLICLVHTAFRV.....KPRPARSAKNAQAELENAY 213

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 405 summaries

Database :

A_Geneseq_0601:*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1165	100.0	213	21	AAV96430 Human TGF-beta bin
2	1164	99.9	213	21	AAV96436 Human TGF-beta bin
3	1157	99.3	213	21	AAV96430 Human TGF-beta bin
4	1157	99.3	213	21	AAV96436 Human TGF-beta bin
5	1157	99.3	213	22	AAV96429 Human TGF-beta bin
6	1136	97.5	213	21	AAV96431 Human TGF-beta bin
7	1064	89.3	213	21	AAV96433 Human TGF-beta bin
8	1021	87.6	211	21	AAV96432 Human TGF-beta bin
9	892.5	76.6	176	21	AAV96434 Human TGF-beta bin
10	351.5	30.2	206	21	AAV96430 Human TGF-beta bin
11					Murine skin cell p

12	351.5	30.2	206	21	AAV76031
13	351.5	30.2	206	22	AAV55920
14	351.5	30.2	206	22	AAV55920
15	349.5	30.0	206	18	AAV09408
16	349.5	30.0	206	19	AAV58704
17	349.5	30.0	206	20	AAV95711
18	349.5	30.0	206	21	AAV10233
19	256	22.0	50	20	AAV12009
20	122.5	10.5	116	18	AAV27654
21	122.5	10.5	116	18	AAV4090
22	117	10.0	23	21	AAV96435
23	108	9.3	168	21	AAV84014
24	104.5	9.0	102	20	AAV03225
25	104.5	9.0	102	21	AAV10277
26	103	8.8	168	21	AAV51132
27	92	7.9	184	19	AAV69293
28	92	7.9	184	20	AAV42173
29	92	7.9	184	21	AAV95961
30	90	7.7	184	20	AAV42172
31	90	7.7	184	20	AAV42174
32	90	7.7	429	20	AAV42175
33	87.5	7.5	184	19	AAV29727
34	87	7.5	246	21	AAV78333
35	87	7.5	391	20	AAV42178
36	87	7.5	558	12	AAV13150
37	86.5	7.4	182	19	AAV29726
38	83.5	7.2	561	12	AAV12342
39	83	7.1	270	19	AAV41250
40	82	7.0	150	21	AAV43879
41	81.5	7.0	1480	13	AAV25079
42	81	7.0	562	6	AAV50342
43	81	7.0	562	9	AAV82582
44	81	7.0	562	11	AAV09288
45	81	7.0	562	11	AAV09290

ALIGNMENTS

RESULT 1	
AAV96430	standard; Protein; 213 AA.
XX	
AC	AAV96430;
XX	
DT	12-SEP-2000 (first entry)
XX	
DE	Human TGF-beta binding protein (BEER) variant V101.
XX	
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW	BEER; variant; V101; gene therapy; antisense therapy; fracture;
KW	chromosome 17q12-21; bone mineralization.
KM	
OS	Homo sapiens.
XX	
XX	
FH	key
FT	Misc-difference 10
FT	/label= V101
FT	/note= "wild type valine has been substituted with Isoleucine"
FT	
FT	
WO200032773-A1.	
XX	
PN	08-JUN-2000.
XX	
PD	24-NOV-1999; 99WO-US27990.
XX	
PR	27-NOV-1998; 98US-0110283.
XX	
XX	(DARM-) DARMIN DISCOVERY LTD.
PA	
XX	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepier BW;
XX	Van Ness J, Winkler DG;
PI	

Murine skin cell p
Skin cell protein,
Human small CCN-11
Human adult retina
Human 5' EST seque
Secreted protein A
Human secreted pro
Mutant human TGF-b
Amino acid sequenc
Human fetal kidney
Murine cerebral ne
Human b57 protein
Mouse gremlin. Ho
Human DRM protein
EGFP/DRM fusion pr
EGFP/DRM fusion pr
DM and b57 protei
Herpes simplex vlr
EGFP/DRM fusion pr
T-PA with -ve char
DM and b57 protei
T-PA with -ve char
Xenopus cerberus p
Human cancer assoc
Drosophila SLIT pr
Human tpa. Homo s
Tissue plasminogen
Sequence of tissue

XX MPI; 2000-412321/35.
DR N-PSDB; AAA29056.
XX
PT Nucleic acids (1) encoding a transforming growth factor beta binding
protein, useful for identifying agents for treating osteopenia,
osteoporosis and fractures
PT
PS Claim 3; Page 119-120; 162pp; English.
XX
CC This shows a variant human transforming growth factor-beta (TGF-beta)
binding protein designated BEER V101, which comprises a substitution of
isoleucine for the wild-type valine at residue 10. The cDNA and protein
CC may be used for prevention, treatment and diagnosis of diseases
CC associated with inappropriate BEER expression. For example, they may be
CC used to treat disorders associated with decreased TGF-beta BP expression.
CC The cDNA or vectors may be administered to treat diseases by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC BEER by expressing inactive proteins or to supplement the patients' own
CC recombinant production of BEER, gene therapy, antisense therapy, as
CC probes for diagnostic assays and for functional studies. BEER may be used
CC to raise antibodies and for identification of BEER modulators. BEER
CC antagonists may be used to increase bone mineral content for the
CC treatment of disorders such as osteopenia, osteoporosis, fractures and
CC other disorders associated with low mineral content.
XX
SQ Sequence 213 AA;
Query Match 100.0%; Score 1165; DB 21; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.1e-105;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOLPLALCTICLVHTAFRVVGGGQAFKNDATETIRLGEYPPPEPLENNKTNRAE 60
Db 1 mqdplalcivcllvhtafvveggwqafkndatellrelgeypppeplennktnrae 60
QY 61 NGRPRPHHFEETKDVSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIG 120
Db 61 nggrpphhfettkdvseyscrelhfttryvtdgpcrsakpvtelvcsgcgparllpnai 120
QY 121 RCKWWRPSPDFRCIPDRRAQVQLCPGGEAPRAKRVLVASCKKRLTRFHNOSELK 180
Db 121 rckwwrpspdfrcipdrayraqvqlcpggeapraikrvlvascckrltrfhngselk 180
QY 181 DEGTAAAPQKGRKPRPARSAKANOAELENAY 213
Db 181 dfgteaarpqkgrkprparsakangaelenay 213
XX
RESULT 2
AA96436
ID AAY96436 standard; Protein: 213 AA.
XX
AC AAY96436;
XX
DT 12-SEP-2000 (first entry)
XX
DE Human TGF-beta binding protein (BEER) variant P38R.
XX
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
BEER; variant; P38R; gene therapy; antisense therapy; fracture;
KW chromosome 17q12-21; bone mineralization.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 38 /label= P38R
FT . /note= "wild type proline has been substituted with
FT arginine"
XX

PN WO200032773-A1.
XX
PD 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
PF
XX 27-NOV-1998; 98US-0110283.
PR
XX
PA (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepier BW;
PI Van Ness J, Winkler DG;
XX
DR MPI; 2000-412321/35.
DR N-PSDB; AAA29062.
XX
PT Nucleic acids (1) encoding a transforming growth factor beta binding
protein, useful for identifying agents for treating osteopenia,
osteoporosis and fractures
PT
PS Disclosure; Page 121; 162pp; English.
XX
CC This shows a variant human transforming growth factor-beta
(TGF-beta) binding protein designated BEER P38R. The encoded protein
CC comprises a substitution of arginine for the wild-type proline at
CC residue 38. The cDNA and protein may be used for prevention, treatment
CC and diagnosis of diseases associated with inappropriate BEER expression.
CC For example, they may be used to treat disorders associated with
CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
CC to treat diseases by rectifying mutations or deletions in a patient's
CC genome that affect the activity of BEER by expressing inactive proteins
CC or to supplement the patients' own production of BEER polypeptides. The
CC nucleic acids may be used for recombinant production of BEER, gene
CC therapy, antisense therapy, as probes for diagnostic assays and for
CC functional studies. BEER may be used to raise antibodies and for
CC identification of BEER modulators. BEER antagonists may be used to
CC increase bone mineral content for the treatment of disorders such as
CC osteopenia, osteoporosis, fractures and other disorders associated with
CC low mineral content.
XX
SQ Sequence 213 AA;
Query Match 99.9%; Score 1164; DB 21; Length 213;
Best Local Similarity 99.5%; Pred. No. 2.7e-105;
Matches 212; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOLPLALCTICLVHTAFRVVGGGQAFKNDATETIRLGEYPPPEPLENNKTNRAE 60
Db 1 mqdplalcivcllvhtafvveggwqafkndatellrelgeypppeplennktnrae 60
QY 61 NGRPRPHHFEETKDVSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIG 120
Db 61 nggrpphhfettkdvseyscrelhfttryvtdgpcrsakpvtelvcsgcgparllpnai 120
QY 121 RCKWWRPSPDFRCIPDRRAQVQLCPGGEAPRAKRVLVASCKKRLTRFHNOSELK 180
Db 121 rckwwrpspdfrcipdrayraqvqlcpggeapraikrvlvascckrltrfhngselk 180
QY 181 DEGTAAAPQKGRKPRPARSAKANOAELENAY 213
Db 181 dfgteaarpqkgrkprparsakangaelenay 213
XX
RESULT 3
AAB26106
ID AAB26106 standard; Protein: 213 AA.
XX
AC AAB26106;
XX
DT 15-JAN-2001 (first entry)
XX
DE Human DAN/Cerberus-related protein 6 (hDRC6) (exons 1 and 4).

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:47:15 ; Search time 47.86 Seconds
(without alignments)
100.150 Million cell updates/sec

Title: US-09-668-021-6
Perfect score: 1165
Sequence: 1 MQLPALCLICLVHPTFRV.....KPRPRASAKANQAELENAY 213

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

```
Database : Issued_Patents_AA:*
1: /usr36/databases/04/iss_AA/COVD.pat
```

```
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCrUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	351.5	30.2	206	4	US-09-188-930-159	Sequence 159, App
2	351.5	30.2	206	4	US-09-188-930-286	Sequence 286, App
3	349.5	30.0	206	1	US-08-668-847B-2	Sequence 2, Appl
4	349.5	30.0	206	1	US-08-668-847B-20	Sequence 20, Appl
5	87.5	7.5	735	3	US-09-191-647-9	Sequence 9, Appl
6	87.5	7.5	735	4	US-09-540-245A-9	Sequence 9, Appl
7	87.5	7.5	735	4	US-09-540-153-9	Sequence 9, Appl
8	82	7.0	270	4	US-08-878-474-1	Sequence 1, Appl
9	81.5	7.0	1480	3	US-09-191-647-7	Sequence 7, Appl
10	81.5	7.0	1480	4	US-09-540-245A-7	Sequence 7, Appl
11	81.5	7.0	1480	4	US-09-540-153-7	Sequence 7, Appl
12	81.5	7.0	1480	5	PCT-US91-09055-2	Sequence 2, Appl
13	80	6.9	510	4	US-08-246-489-2	Sequence 2, Appl
14	79	6.8	336	4	US-08-804-166-8	Sequence 8, Appl
15	79	6.8	336	4	US-08-910-991-8	Sequence 8, Appl
16	79	6.8	566	1	US-08-073-383-4	Sequence 4, Appl
17	79	6.8	566	5	PCT-US94-06365-4	Sequence 4, Appl
18	78	6.7	1210	2	US-08-484-438-7	Sequence 7, Appl
19	78	6.7	1210	2	US-08-475-035-7	Sequence 4, Appl
20	77.5	6.7	76	5	PCT-US91-09055-9	Sequence 9, Appl
21	76	6.5	546	6	5200340-6	Patent No. 5200340
22	76	6.5	562	2	US-08-560-098A-50	Sequence 50, Appl
23	76	6.5	562	2	US-08-883-795A-38	Sequence 38, Appl
24	76	6.5	562	6	5185259-3	Patent No. 5185259
25	76	6.5	562	6	5200340-2	Patent No. 5200340
26	76	6.5	562	6	5344773-2	Patent No. 5344777
27	74	6.4	470	2	US-08-946-241B-2	Sequence 2, Appl

28	74	6.4	470	3	US-09-309-053-2	Sequence 2, Appl
29	74	6.4	479	2	US-08-946-241B-9	Sequence 9, Appl
30	74	6.4	479	3	US-09-309-053-9	Sequence 9, Appl
31	73.5	6.3	301	4	US-08-303-861-21	Sequence 21, Appl
32	73.5	6.3	301	4	US-09-011-073A-1	Sequence 1, Appl
33	73.5	6.3	685	2	US-08-878-989-1	Sequence 1, Appl
34	73.5	6.3	685	3	US-09-135-282-2	Sequence 1, Appl
35	73.5	6.3	685	4	US-09-272-796-1	Sequence 1, Appl
36	73.5	6.3	685	4	US-09-505-744-2	Sequence 2, Appl
37	73	6.2	301	4	US-09-230-421-2	Sequence 2, Appl
38	72.5	6.2	375	1	US-08-468-847B-13	Sequence 13, Appl
39	72	6.1	379	1	US-08-946-847B-11	Sequence 11, Appl
40	71.5	6.1	521	3	US-08-948-564-14	Sequence 14, Appl
41	71	6.1	449	1	US-08-103-942A-4	Sequence 4, Appl
42	71	6.1	1259	1	US-09-187-049-13	Sequence 13, Appl
43	70.5	6.1	390	1	US-08-347-792-15	Sequence 15, Appl
44	70.5	6.1	390	1	US-08-431-357-15	Sequence 15, Appl
45	70.5	6.1	390	4	US-08-392-542-5	Sequence 15, Appl

ALIGNMENTS

```

RESULT 1
US-09-188-930-159
: Sequence 159, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Stiechan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Onrust, Rene
: TITLE OF INVENTION: Compositions Isolated
: FILE OF INVENTION: and Methods For Their
: FILE REFERENCE: 11000.101c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 159
: LENGTH: 206
: TYPE: RNT
: ORGANISM: mouse
US-09-188-930-159

```

Query Match	30.28;	Score 351.5;	DB 4;	Length 206;
Best Local Similarity	40.48;	Pred. No. 5.9e-30;		
Matches 82;	Conservative 32;	Mismatches 68;	Indels 21;	Gaps 8

[illegible]

RESULT 2
US-09-188-930-286
; Sequence 286, Application US/09188930R
; Patent No. 6150502
; GENERAL INFORMATION:

APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011C1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 286
LENGTH: 206
TYPE: PR1
ORGANISM: Mouse
US-09-188-930-286

Query Match 30.2%; Score 351.5; DB 4; Length 206;
Best Local Similarity 40.4%; Pred. No. 5.9e-30;
Matches 82; Conservative 32; Mismatches 68; Indels 21; Gaps 8;

9 LCLICLHVFARVYEGOGWAFKNDATETIRELGEYPEPPPELENNKTWNRANGCRPPHH 68
13 LCLICLHVFARVYEGOGWAFKNDATETIRELGEYPEPPPELENNKTWNRANGCRPPHH 68
69 PFTKDV--EYSCRELHFTRYVDGPCRSAPVTELVCSGQCGPARLLPNAI 122
61 SSTGLNRSSQWRCVNDKRTQRIOLQCGD-STRTYKITYVTACKCKRYTQHSHSNF 120
123 KMW-RPSGPDPCIPDRYRAORVCLPGGEAPRARKVRLVASCCKRLTFHNSQELKD 181
121 KYMSRRSSQWRCVNDKRTQRIOLQCGD-STRTYKITYVTACKCKRYTQHSHSNF 179
182 FGTETARPQKGRKPPRA-RSAK 203
180 ESVSPAKPAQHRRKRASKSK 202

RESULT 3
US-08-468-847B-2
Sequence 2, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-2

Query Match 30.0%; Score 349.5; DB 1; Length 206;
Best Local Similarity 40.7%; Pred. No. 9.7e-30;
Matches 85; Conservative 32; Mismatches 67; Indels 25; Gaps 9;

3 LPLALCLICLHVFARVYEGOGWAFKNDATETIRELGEYPEPPPELENNKTWNRANG 62
11 LPLA-----CLIKSCL-----AFKNDATETILSHVVKP-VPAHPSNSTLNQARNG 56
63 GRPHHPETKDV--EYSCRELHFTRYVDGPCRSAPVTELVCSGQCGPARLLPNAI 119
57 GR--HFSNTGLDRNTRYOVGCRELSTKYISDQCTSISPLKELVCAGECLPLPLPMWI 114
120 GRG--KMW-RPSGPDPCIPDRYRAORVCLPGGEAPRARKVRLVASCCKRLTFHNF 175
115 GGGYGTKYMSRRSSQWRCVNDKRTQRIOLQCGD-STRTYKITYVTACKCKRYTQHSH 173
176 OSELKDEGTETARPQKGRKPPRA-RSAK 203
174 ESHNFSMSPAKPVQHHRRKRASKSK 202

RESULT 4
US-08-468-847B-2
Sequence 20, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:49:44 ; Search time 54.26 Seconds

(without alignments)
299,027 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 1169

Sequence: 1 MQLPLALCVLCLVHAFRV.....KPRPARCANKANQALENAY 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Archived: 219241 segs, 76174552 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR,68:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	268.5	23.0	134	2 T08710	hypothetical prote
2	99.5	8.5	159	2 I51373	luteinizing hormon
3	94.5	8.1	158	2 A61091	lutropin beta chai
4	89	7.6	118	2 PM0141	lutropin beta chai
5	88.5	7.6	601	2 T22025	hypothetical prote
6	87	7.4	118	2 PM0139	lutropin beta chai
7	84	7.2	1469	2 B36655	slit protein 2 pre
8	83.5	7.1	166	2 I51242	luteinizing hormon
9	83	7.1	2531	2 T16743	hypothetical prote
10	81.5	7.0	969	2 A70912	probable leus prot
11	81.5	7.0	1480	2 A36655	slit protein 1 pre
12	81	6.9	562	2 UKHUT	t-plasminogen acti
13	80	6.8	1138	2 G85077	probable polyprote
14	79.5	6.8	165	1 KPHUB	chorio gonadotropin
15	79.5	6.8	824	2 T10020	leucine--trna 11ga
16	79.5	6.8	972	2 T10023	leucine--trna 11ga
17	79.5	6.8	2333	1 GNNY2F	genome polyprotein
18	79	6.8	270	2 S71793	head-inducing fact
19	78.5	6.7	128	2 S74085	lutropin beta chai
20	78.5	6.7	1042	2 A57534	mucin 5AC (clone L
21	78	6.6	510	2 A42050	insulinoma-associa
22	77	6.6	343	2 G35070	apolipoprotein H-r
23	76.5	6.5	145	2 I37231	beta-gonadotropin
24	76.5	6.5	356	1 VCVMB6	env polyprotein pr
25	76.5	6.5	1223	2 S29717	adenylate cyclase
26	76.5	6.5	2098	2 T18397	protein CRP - mal
27	75	6.4	636	1 VCVMS	env polyprotein
28	75	6.4	864	1 H85335	hypothetical prote
29	75	6.4	864	2 T04518	hypothetical prote

30	75	6.4	1317	2 T03748	apoptosis associat
31	74.5	6.4	682	2 A44493	serum-inducible ki
32	74.5	6.4	690	2 T33321	hypothetical prote
33	74.5	6.4	1224	2 T00059	hypothetical prote
34	74	6.4	166	2 T39546	hypothetical prote
35	74	6.3	269	2 B69381	hypothetical prote
36	74	6.3	328	2 T151215	inhibin alpha-subu
37	74	6.3	367	2 T18839	hypothetical prote
38	74	6.3	381	2 S60561	I kappa B-like pro
39	74	6.3	383	2 A23516	Balbani ring 1 ch
40	74	6.3	398	2 A35281	intergenetary muc
41	74	6.3	405	2 S33601	cell adhesion mole
42	74	6.3	992	2 A31666	hypothetical prote
43	74	6.3	1589	2 T42233	sumaxillary mucin
44	74	6.3	2153	2 T30074	hypothetical prote
45	73.5	6.3	221	2 S70009	glutamate/proline-

ALIGNMENTS

RESULT 1
T08710
hypothetical protein DKFZP564D206.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
C/Accession: T08710
R/Wambull, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, March 1999
A/Reference number: 216471
A/Accession: T08710
A/Molecule type: mRNA
A/Residues: 1-134 <MAN>
A/Cross-references: EMBL:AL050024
A/Experimental source: fetal brain; clone DKFZP564D206
C/Genetics:
A/Note: DKFZP564D206.1

Query Match 23.0%; Score 268.5; DB 2; Length 134;
Best Local Similarity 44.4%; Pred. No. 1; Je-16;

Matches 55; Conservative 21; Mismatches 43; Indels 5; Gaps 3;

OY 80 CRELHFRYVYDGPCKSAKPYTELVCSGGCPARLPLNAIGRG---KMW-RPSGDPFCIT 135

Db 3 CRELHFRYVYDGPCKSAKPYTELVCSGGCPARLPLNAIGRG---KMW-RPSGDPFCIT 62

OY 136 PDYRAQRVQLCPGGAAPRAKRVLVASCKRRLTFHNSQELKDPGEAARPOKGRK 195

Db 63 NDKTRQRIQLCQDQGR-RTYKIIVYVACKCKRYTROHNESSHNEFSMPAKRYQHRE 121

OY 196 RPKA 199

Db 122 RKRA 125

RESULT 2

I51373

C/Species: Melospiza gallopavo (common turkey)

C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999

C/Accession: I51373

R/You, S.; Foster, L.R.; Salsby, J.L.; el Halawani, M.E.; Foster, D.N.

A/Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonado

A/Reference number: I51373; MUID:95290073

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-159 <YON>

A/Cross-references: GB:I55519; NID:g530952; PTDN:AAA14125.1; PTD:g530953

C/Genetics:

A/Gene: LH-beta

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:13 ; Search time 31.13 Seconds

(Without alignments)
234.385 Million cell updates/sec

Title: US-09-668-021-10

Sequence: 1 MOLPLALCVCLVHAHFV.....KPPRARGAKNAQAELENAY 213

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

rchd: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99.5	8.5	159	1	LSHB_MELGA
2	89	7.6	118	1	LSHB_PHYCA
3	87	7.4	118	1	LSHB_BALCA
4	86.5	7.4	870	1	BCAL_HUMAN
5	85.5	7.3	141	1	LSHB_TRLVU
6	83.5	7.1	166	1	LSHB_CORJA
7	81.5	7.0	969	1	SLIT_MYCTU
8	81.5	7.0	1480	1	SLIT_DROME
9	81	6.9	562	1	TPA_HUMAN
10	79.5	6.8	165	1	CGHB_HUMAN
11	79.5	6.8	958	1	HIG_DROME
12	79.5	6.8	972	1	STL_MYCLE
13	79.5	6.8	2333	1	POLG_FMYV1
14	79	6.8	180	1	DNA_HUMAN
15	79	6.8	470	1	KLF4_HUMAN
16	79	6.8	474	1	KLF4_MOUSE
17	78.5	6.7	128	1	LSHB_STRCA
18	78	6.7	510	1	IA1_HUMAN
19	77.5	6.6	318	1	HR25_MOUSE
20	77	6.6	141	1	LSHB_CERS1
21	77	6.6	355	1	KLF4_HUMAN
22	77	6.6	625	1	DUS8_HUMAN
23	76.5	6.5	336	1	ENV_FRSTB
24	76.5	6.5	951	1	STR8_HUMAN
25	76	6.5	138	1	LSHB_MACRU
26	76	6.5	1233	1	NME3_HUMAN
27	75.5	6.5	755	1	RRE1_HUMAN
28	75	6.4	636	1	ENV_MCFP
29	74.5	6.4	682	1	SNK_MOUSE
30	74.5	6.4	682	1	SNK_MOUSE
31	74.5	6.4	685	1	SNK_MOUSE
32	74	6.3	398	1	MOB1_XENLA
33	74	6.3	968	1	BCAL_RAT

34	74	6.3	992	1	EBN6_EBV
35	73.5	6.3	143	1	LSHB_FELCA
36	73.5	6.3	409	1	ENV2_FRSEV
37	73.5	6.3	1056	1	MUC5_HUMAN
38	73.5	6.3	2142	1	BAT2_HUMAN
39	73	6.2	191	1	CBX5_HUMAN
40	73	6.2	301	1	UT49_HSV21
41	73	6.2	619	1	BCHD_CHLVI
42	72.5	6.2	318	1	HR25_RAT
43	72.5	6.2	450	1	A2AA_MOUSE
44	72.5	6.2	1157	1	Y182_HUMAN
45	72	6.2	174	1	BAR1_CHITE

ALIGNMENTS

RESULT 1	
LSHB_MELGA	
AC P45646;	STANDARD; PRT; 159 AA.
DT 01-NOV-1995 (Rel. 32, Created)	
DT 01-NOV-1995 (Rel. 32, Last sequence update)	
DT 01-OCT-2000 (Rel. 40, Last annotation update)	
DE LUTROPIN BETA CHAIN PRECURSOR (LUTRINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).	
GN LHB.	
OS Meleagris gallopavo (Common turkey).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.	
OX NCBI_TaxID=9103;	
RP [1]	
RC SEQUENCE FROM N.A.	
RC TRISSUE-Pituitary;	
RA MEDLINE=95290073; PubMed=7772235;	
RA You S., Foster L.K., Sillsby J.L., el Halawani M.E., Foster D.N.;	
RT "Sequence analysis of the turkey LH beta subunit and its regulation by gonadotropin-releasing hormone and prolactin in cultured pituitary cells.";	
RT J. Mol. Endocrinol. 14:117-129(1995).	
CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.	
CC -!- SUBUNIT: HETRODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONTERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,	
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.	
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC -----	
CC EMBL: L35519; AAA74125.1; ALT_INIT.	
CC HSSP: P01233; IHRP.	
CC InterPro: IPR000359; -	
CC InterPro: IPR001545; -	
CC InterPro: IPR002400; -	
CC Pfam: PF00007; Cys_knot.1.	
CC PRINTS: PR00438; GFCISKNOT.	
CC PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.	
CC PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.	
CC Hormone; Signal; Glycoprotein.	
FT SIGNAL 1 39	POTENTIAL.
FT CHAIN 40 159	LUTROPIN BETA CHAIN.
FT DISULFID 48 96	BY SIMILARITY.
FT DISULFID 62 111	BY SIMILARITY.
FT DISULFID 65 149	BY SIMILARITY.
FT DISULFID 73 127	BY SIMILARITY.
FT DISULFID 77 129	BY SIMILARITY.


```

RESULT 2
Q9PWB0 PRELIMINARY; PRT; 272 AA.
AC Q9PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CEREBRUS HOMOLOG.
GN CER.
OS Gallus gallus (Chicken).
OC Archyotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
ON NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9439862; Pubmed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RT "Cerebus regulates left-right asymmetry of the embryonic head and
heart."
Curr. Biol. 9:931-938(1999).
EMBL: AF139721; AAD51610.1; -.
InterPro: IPR000359; -.
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT; 1.
SQ SMOU0041; CT; 1.
SEQUENCE 272 AA; 36E4C9F719711BCA CRC64;

Query Match
Best Local Similarity 9.2%; Score 107; DB 13; Length 272;
Matches 32; Conservative 21; Mismatches 56; Indels 26; Gaps 5;

QY 45 EPPLENNK-----TNNRAENGRRPHHPETKDVSEYSCRELHRTYVTDGPRCSAKPV 100
DB 129 EPPYRKDAKKFMDHFMRLKNSASEVYLPKTNEMHQETCRTPFSQVAHSCERK-IV 187
QY 101 TELVSCGCCPARLIPNAIGRKWMRSPGDFR-----CIPDVRRAQVOLLCPGGA 153
DB 188 QNNLCFGKCSS-----FHVPGPDRLYTFCKCLPTKFSMKHDLNCT-SSV 233
QY 154 PRARKRLVASCKK 168
DB 234 PVYKKWIVEECNCE 248

RESULT 3
Q9PUK2 PRELIMINARY; PRT; 272 AA.
AC Q9PUK2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CARONTE.
GN CAR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archyotia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
ON NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX Rodriguez Esteban C., Capdevilla J., Economides A.N., Pascual J.,
RA Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-like protein, mediates the establishment of
embryonic left-right asymmetry."
RL Nature 0:0-0(1999).
EMBL: AF179484; AAD5581.1; -.
InterPro: IPR000359; -.
DR InterPro: IPR001839; -.
DR ProDom: PD000357; -.
DR PROSITE: PS01225; CTCK_2; 1.

```

```

DR SMART: SM00041; CT; 1.
SQ SMOU0041; CT; 1.
SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match
Best Local Similarity 9.0%; Score 105; DB 13; Length 272;
Matches 32; Conservative 20; Mismatches 57; Indels 26; Gaps 5;

QY 45 EPPLENNK-----TNNRAENGRRPHHPETKDVSEYSCRELHRTYVTDGPRCSAKPV 100
DB 129 EPPYRKDAKKFMDHFMRLKNSASEVYLPKTNEMHQETCRTPFSQVAHSCERK-IV 187
QY 101 TELVSCGCCPARLIPNAIGRKWMRSPGDFR-----CIPDVRRAQVOLLCPGGA 153
DB 188 QNNLCFGKCSS-----FHVPGPDRLYTFCKCLPTKFSMKHDLNCT-SSV 233
QY 154 PRARKRLVASCKK 168
DB 234 PVYKKWIVEECNCE 248

RESULT 4
Q9H772 PRELIMINARY; PRT; 168 AA.
AC Q9H772;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CDNA: FLJ21195 FIS, CLONE COL00185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX KAWABATA A., HIKIJI T., KODATAKE N., INAGAKI H., IKEMA Y., OKAMOTO S.,
RA KAWABATA A., HIKIJI T., KODATAKE N., INAGAKI H., IKEMA Y., OKAMOTO S.,
RA Okitani T., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
EMBL: AK024648; BAB15026.1; -.
DR AK024648; BAB15026.1; -.
SEQUENCE 168 AA; 19320 MW; D5A4E4EB18BF8C0E CRC64;

Query Match
Best Local Similarity 8.5%; Score 99; DB 4; Length 168;
Matches 45; Conservative 25; Mismatches 66; Indels 48; Gaps 9;

QY 2 QLPALCLVCLVHAARVVEGGQWAFKNDATETIIPELGEYPEPPLENNKTNRAEN 61
DB 4 KLSLSEFLVAVLVKA-----EARKNRPA-----GALSPYKDGSSN-----N 41
QY 62 GGRPHHPETKDVSEYSCRELHRTYVTDGPRCSAKPVTELVCSCQ 108
DB 42 SERHQIQIKFVLAISQDALVYTERKLYKSDCTQPLRQVSEEGRS-RTLNRFCTGQ 100
QY 109 CGPARLPNAIGRKWMRSPGDFR-----CIPDVRRAQVOLLCPGGAAP-PRARKRLVA 163
DB 101 CN-SFYPRHV-----KKEESFQSCAFCKPQRYVSVLVELECGDLPPLKRIQKVK 153
QY 164 SCCK 167
DB 154 QCRG 157

RESULT 5
Q9UEM9 PRELIMINARY; PRT; 493 AA.
AC Q9UEM9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:42 ; Search time 80.28 Seconds
(without alignments)
160.848 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 1169

Sequence: 1 MQLPLALICVLCVLAHAFRV.....KPRPRARAKANQAELENAV 213

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Residues: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_0601.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1169	100.0	213	21	AA196431
2	1146	98.0	213	21	AA196431
3	1146	98.0	213	21	AA196431
4	1146	98.0	213	22	AA197589
5	1137	97.3	213	21	AA196436
6	1136	97.2	213	21	AA196430
7	1073	90.0	213	21	AA196433
8	1052	90.0	211	21	AA196432
9	1051	89.9	367	21	AA196432
10	920.5	78.7	176	21	AA196434
11	348	29.8	206	18	AAW09408

12	348	29.8	206	19	AAW58704	Human small CCN-11
13	348	29.8	206	20	AAW95711	Homo sapiens fetal
14	348	29.8	206	21	AA196431	Human adult retina
15	346	29.6	206	21	AA197581	Murine skin cell p
16	346	29.6	206	21	AA196431	Murine skin cell p
17	346	29.6	206	22	AA196431	Skin cell protein,
18	346	29.6	206	22	AA196431	Skin cell protein,
19	261	22.3	50	20	AA12009	Human 5' EST sec
20	122.5	10.5	116	18	AAW27654	Secreted protein A
21	122.5	10.5	116	18	AAW44090	Human secreted pro
22	113	9.7	23	21	AA196435	Mutant human TGF- β
23	103.5	8.9	102	20	AA196435	Amino acid sequenc
24	103.5	8.9	102	21	AA196435	Human fetal kidney
25	99	8.5	168	21	AA196435	Amino acid sequenc
26	94	8.0	168	21	AA196435	Murine cerebral ne
27	93	8.0	712	21	AA196435	Human ORFX ORF2911
28	92	7.9	558	12	AA196435	T-PA with -ve char
29	89.5	7.7	1428	21	AA196435	Caspase 8-interact
30	89	7.6	336	18	AAW33360	TBP(20-190)/hcg-be
31	88.5	7.6	561	12	AA196435	T-PA with -ve char
32	88	7.5	184	19	AAW69293	Human b57 protein
33	88	7.5	184	20	AA196435	Mouse DM protein
34	88	7.5	184	21	AA196435	Human gremilin. Ho
35	86.5	7.4	870	21	AA196435	Human ORFX ORF1506
36	86	7.4	562	6	AA196435	Human tpa. Homo s
37	86	7.4	562	9	AA196435	Tissue plasminogen
38	86	7.4	562	11	AA196435	Sequence of tissue
39	86	7.4	562	11	AA196435	Sequence of tissue
40	85.5	7.3	145	12	AA196435	hcg/ELH chimera, E
41	85.5	7.3	145	12	AA196435	hcg/ELH chimera, A
42	85	7.3	562	11	AA196435	Sequence of tissue
43	84.5	7.2	184	19	AAW29727	DAN and b57 protei
44	84.5	7.2	184	20	AA196435	Human DM protein se
45	84.5	7.2	184	20	AA196435	Rat DRM protein se

ALIGNMENTS

RESULT 1	AA196431	standard; Protein; 213 AA.
XX	AA196431	
AC	AA196431	
DT	12-SEP-2000	(first entry)
XX		
DE	Vervet TGF-beta binding protein (BEER).	
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KM	BEER; gene therapy; antisense therapy; fracture; bone mineralization.	
XX		
OS	Cercopithecus pygerythrus.	
XX		
PN	WO200032773-A1.	
PD	08-JUN-2000.	
XX		
PF	24-NOV-1999.	99WO-US27990.
XX		
PR	27-NOV-1998.	98US-0110283.
PA	(DARW-) DARWIN DISCOVERY LTD.	
XX		
PI	Brunow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
XX	Van Ness J, Winkler DG;	
DR	WPI; 2000-412321/35.	
XX	N-PSDB; AAA29057.	
PT	Nucleic acids (I) encoding a transforming growth factor beta binding	
PT	protein, useful for identifying agents for treating osteopenia,	
PT	osteoporosis and fractures	

```
XX Claim 4; Page 122-123; 162pp; English.
PS
XX This shows a verret transforming growth factor-beta (TGF-beta)
CC binding protein designated vBER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BERR expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BERR by
CC expressing inactive proteins or to supplement the patient's own production
CC of BERR polypeptides. The nucleic acids may be used for recombinant
CC production of BERR, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BERR may be used to raise
CC antibodies and for identification of BERR modulators. BERR antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.
XX
SQ Sequence 213 AA;

Query Match 100.0%; Score 1169; DB 21; Length 213;
Best Local Similarity 100.0%; Pred. No. 9e-106;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOLPLALCLVCLVHAARFVVEGCGMAFKNDATETIPELGEPPEPELENKTMNRAE 60
   |||||||
DB 1 mqlplalcivcllvhaafvveggwqafkndateilpelgeypppepeleennktmnae 60
QY 61 NGRPRHPHPEETKDVSEYSCRELHFTFRVTDGPCRSAPVTETVCSGCGPARLLPNAIG 120
   |||||||
DB 61 nggrprhphpeetkdvseyscrelhftfrvtdgpcrsakpvteivcsqcgparllpnaig 120
QY 121 RKGMMWRPSGDPFCIPDRYRAORVOLLCPGGAAPRARKVRLVASCKCKRLTFPHNOSLTK 180
   |||||||
DB 121 rgkwmwrpsgdpfcipdryraqrvqlldpggaaprarvrlvasckckrltrfhnqselk 180
QY 181 DFGPEAARPOKGRKPRPARAGAKANAQAELENAY 213
   |||||||
DB 181 dfgpeaarpqkgrkprparagakanqaelenay 213

RESULT 2
AAB26106
ID AAB26106 standard; Protein: 213 AA.
XX
AC AAB26106;
XX
XX 15-JAN-2001 (first entry)
DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
XX
XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
KW antagonist; BMP; cell growth; cell differentiation; bone formation;
KW gene therapy.
XX
OS Homo sapiens.
XX
XX WO200055193-A2.
XX
XX 21-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-US05537.
XX
XX 12-MAR-1999; 99US-0124118.
XX
XX (REGF-) REGENERON PHARM INC.
XX
XX Economides AN;
XX
XX WPI; 2000-638179/61.
XX
XX N-PSDB; AAA94051.
XX
DR
```

```
XX Novel isolated, human DNA/Cerberus related protein 6 which include
PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
PT acids encoding the proteins which are useful as probes and primers
XX
XX Claim 8; Fig 3; 40pp; English.
XX
XX The present sequence comprises the amino acid sequence encoded by exons 1
CC and 4 of the human DAN/cerberus-related protein 6 (hDCR6) coding
CC sequence. The coding sequence was isolated from a human kidney cDNA
CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
CC related to the DAN and DCR5 proteins, both of which act as antagonists of
CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
CC protein can be used as immunogens, modulators of cell function, growth
CC and differentiation, to reduce undesirable bone formation, to identify
CC DCR6 binding agents, in diagnosis, and in gene therapy.
XX
SQ Sequence 213 AA;

Query Match 98.0%; Score 1146; DB 21; Length 213;
Best Local Similarity 98.1%; Pred. No. 1.5e-103;
Matches 209; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MOLPLALCLVCLVHAARFVVEGCGMAFKNDATETIPELGEPPEPELENKTMNRAE 60
   |||||||
DB 1 mqlplalcivcllvhaafvveggwqafkndateilpelgeypppepeleennktmnae 60
QY 61 NGRPRHPHPEETKDVSEYSCRELHFTFRVTDGPCRSAPVTETVCSGCGPARLLPNAIG 120
   |||||||
DB 61 nggrprhphpeetkdvseyscrelhftfrvtdgpcrsakpvteivcsqcgparllpnaig 120
QY 121 RKGMMWRPSGDPFCIPDRYRAORVOLLCPGGAAPRARKVRLVASCKCKRLTFPHNOSLTK 180
   |||||||
DB 121 rgkwmwrpsgdpfcipdryraqrvqlldpggaaprarvrlvasckckrltrfhnqselk 180
QY 181 DFGPEAARPOKGRKPRPARAGAKANAQAELENAY 213
   |||||||
DB 181 dfgpeaarpqkgrkprparagakanqaelenay 213

RESULT 3
AAI96429
ID AAY96429 standard; Protein: 213 AA.
XX
AC AAY96429;
XX
XX 12-SEP-2000 (first entry)
DE Human TGF-beta binding protein (BERR).
XX
XX Human TGF-beta binding protein (BERR).
DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BERR; gene therapy; antisense therapy; fracture; bone mineralization.
XX
OS Homo sapiens.
XX
XX WO200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX
XX (DARW-) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
XX Van Ness J, Winkler DG;
XX
XX WPI; 2000-412321/35.
XX
XX N-PSDB; AAA29055.
XX
XX Nucleic acids (I) encoding a transforming growth factor beta binding
PT
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:47:16 ; Search time 47.86 Seconds
(without alignments)
100.150 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 1169

Sequence: 1 MQLPLALCLVCLLVHAAFRV.....KPRPRARAKAQAQLENNAY 213

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5
rched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2.6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2.6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2.6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2.6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2.6/ptodata/2/1aa/PCBUS.COMB.pep:*
6: /cgn2.6/ptodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	348	29.8	206	1	US-08-468-847B-2
2	348	29.8	206	1	US-08-468-847B-20
3	346	29.6	206	4	US-09-188-930-159
4	346	29.6	206	4	US-09-188-930-286
5	89	7.6	336	4	US-08-804-166-8
6	89	7.6	336	4	US-08-910-991-8
7	88.5	7.6	735	3	US-09-191-647-9
8	88.5	7.6	735	4	US-09-540-245A-9
9	88.5	7.6	735	4	US-09-540-153-9
10	81.5	7.0	1480	3	US-09-191-647-7
11	81.5	7.0	1480	4	US-09-540-245A-7
12	81.5	7.0	1480	4	US-09-540-153-7
13	81.5	7.0	1480	5	PCT-US91-09055-2
14	81	6.9	545	6	5200340-6
15	81	6.9	562	2	US-08-560-098A-50
16	81	6.9	562	2	US-08-883-795A-38
17	81	6.9	562	6	5185259-3
18	81	6.9	562	6	5200340-2
19	81	6.9	562	6	5344773-2
20	80	6.8	307	4	US-08-804-166-4
21	80	6.8	307	4	US-08-910-991-4
22	79.5	6.8	145	1	US-08-298-189B-1
23	79.5	6.8	145	1	US-08-475-213-10
24	79.5	6.8	145	1	US-08-395-238-2
25	79.5	6.8	145	4	US-09-142-320-12
26	79.5	6.8	145	4	US-09-142-320-13
27	79.5	6.8	145	4	US-09-142-320-14

28	79.5	6.8	145	4	US-09-142-320-15	Sequence 15, Appl
29	79.5	6.8	145	4	US-08-918-288-68	Sequence 68, Appl
30	79.5	6.8	145	4	US-09-282-357-68	Sequence 68, Appl
31	79.5	6.8	265	4	US-08-918-288-3	Sequence 3, Appl
32	79.5	6.8	265	4	US-08-918-288-39	Sequence 39, Appl
33	79.5	6.8	265	4	US-09-282-357-3	Sequence 3, Appl
34	79.5	6.8	265	4	US-09-282-357-39	Sequence 39, Appl
35	79	6.8	270	2	US-08-878-474-1	Sequence 2, Appl
36	79	6.8	470	2	US-08-946-241B-2	Sequence 2, Appl
37	79	6.8	470	2	US-09-309-053-2	Sequence 2, Appl
38	79	6.8	479	2	US-08-946-241B-9	Sequence 9, Appl
39	79	6.8	479	3	US-09-309-053-9	Sequence 9, Appl
40	78.5	6.7	165	2	US-08-709-924-2	Sequence 2, Appl
41	78.5	6.7	165	2	US-08-709-925-2	Sequence 2, Appl
42	78.5	6.7	181	4	US-08-918-288-36	Sequence 36, Appl
43	78.5	6.7	181	4	US-09-282-357-36	Sequence 36, Appl
44	78	6.7	510	4	US-08-246-489-2	Sequence 2, Appl
45	77.5	6.6	145	1	US-08-425-673-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-468-847B-2
Sequence 2, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-2

Query Match 29.8%; Score 348; DB 1; Length 206;
Best Local Similarity 40.5%; Pred. No. 3.1e-29;
Matches 83; Conservative 31; Mismatches 65; Indels 26; Gaps 9;
QY 3 LPALCLVCLLVHAAFRVVGSGQWAFKNDATRIIFELGEYEP-PPLELNKTKNNRAEN 61

```

DB 11 LPLA-----CILMKSCLE-----AFKNDATEIL--YSHVVKPVPAHSSNSTLNOARN 55
QY 62 GGRPHHPETKDV---EYSCRELHFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNA 118
DB 56 GGR--HFSNTGLDRNTRVOYGCRELSTKTYISDSGCTISPLKELVCAECCLPLPVLPMW 113
QY 119 IGRG---KWM-RPSGDPFCIPDRYRAORVOLLCPGGAAPRAKRVLVASCCKRLTRFH 174
DB 114 IGGYGTKTYMSRRSSQEMRCVNDKTRTORIQLOCDGST-RTYKITVYVACKCKRYTROH 172
QY 175 NOSELKDFGEAPARPOKGRKPRPA 199
DB 173 NESHNFESMSPAKPVQHHRRKRA 197

```

RESULT 2

US-08-468-847B-20
Sequence 20, Application US/08468847B
Patent No. 5780263

GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-20

Query Match 29.8%; Score 348; DB 1; Length 206;
Best Local Similarity 40.5%; Pred. No. 3,1e-29;
Matches 83; Conservative 31; Mismatches 65; Indels 26; Gaps 9;

```

QY 3 LPLALVCLVLAHFAFRVVGQGMQAFKNDATEILPELGEYEP-PPLELNKKTNNRAEN 61
DB 11 LPLA-----CILMKSCLE-----AFKNDATEIL--YSHVVKPVPAHSSNSTLNOARN 55
QY 62 GGRPHHPETKDV---EYSCRELHFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNA 118
DB 56 GGR--HFSNTGLDRNTRVOYGCRELSTKTYISDSGCTISPLKELVCAECCLPLPVLPMW 113

```

```

QY 119 IGRG---KWM-RPSGDPFCIPDRYRAORVOLLCPGGAAPRAKRVLVASCCKRLTRFH 174
DB 114 IGGYGTKTYMSRRSSQEMRCVNDKTRTORIQLOCDGST-RTYKITVYVACKCKRYTROH 172
QY 175 NOSELKDFGEAPARPOKGRKPRPA 199
DB 173 NESHNFESMSPAKPVQHHRRKRA 197

```

RESULT 3

US-09-188-930-159
Sequence 159, Application US/09188930A
Patent No. 6150502

GENERAL INFORMATION:

APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated from Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 159
LENGTH: 206
TYPE: PRT
ORGANISM: mouse
US-09-188-930-159

Query Match 29.6%; Score 346; DB 4; Length 206;
Best Local Similarity 40.2%; Pred. No. 5,1e-29;
Matches 80; Conservative 31; Mismatches 66; Indels 22; Gaps 8;

```

QY 9 LVCLLVHAHFAFRVVGQGMQAFKNDATEILPELGEYEP-PPLELNKKTNNRAENGRPH 67
DB 13 LCLTLMNCL-----AFKNDATEIL--YSHVVKPVPAHSSNSTLNOANGR--H 59
QY 68 HPEFTKDV---EYSCRELHFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIRG-- 122
DB 60 FSTGLDRNTRVOYGCRELSTKTYISDSGCTISPLKELVCAECCLPLPVLPMWIGGYG 119
QY 123 -KWM-RPSGDPFCIPDRYRAORVOLLCPGGAAPRAKRVLVASCCKRLTRFHNOSELK 180
DB 120 TKYMSRRSSQEMRCVNDKTRTORIQLOCDGST-RTYKITVYVACKCKRYTROHNESSHN 178
QY 181 DFGPEAPARPOKGRKPRPA 199
DB 179 FEVSYPAPKPVQHHRRKRA 197

```

RESULT 4

US-09-188-930-286
Sequence 286, Application US/09188930A
Patent No. 6150502

GENERAL INFORMATION:

APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated from Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 286
LENGTH: 206

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:49:45 ; Search time 54.26 Seconds
(without alignments) 296.219 Million cell updates/sec

Title: US-09-668-021-12

Perfect score: 1171

Sequence: 1 M0PSLAPCLICLIVHAFA.....KPPRGARGAQAELENAY 211

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

riched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249.5	21.3	134	2 T08710	hypothetical prote
2	93	7.9	118	2 PN0141	lutropin beta chai
3	92.5	7.9	159	2 I51373	lutreinizng hormon
4	91	7.8	118	2 PN0139	lutropin beta chai
5	89.5	7.6	158	2 A61091	lutropin beta chai
6	88	7.5	332	2 S46947	ribosomal protein
7	84.5	7.2	1223	2 S29717	adenylate cyclase
8	82.5	7.0	334	2 T04572	hypothetical prote
9	82	7.0	301	1 WMBEP9	ul49 protein - hum
10	81.5	7.0	349	2 S49606	ribosomal protein
11	81.5	7.0	2182	2 T14320	calcineurin inhibi
12	81.5	7.0	13288	2 T03099	mucin, submaxillar
13	81	6.9	405	2 S33601	cell adhesion mole
14	80.5	6.9	270	2 S71793	head-inducing fact
15	80.5	6.9	348	2 S75548	staloglycoproteina
16	80.5	6.9	975	2 S33121	homeotic protein C
17	80	6.8	866	2 T29197	hypothetical prote
18	80	6.8	888	2 T58378	tyrosine kinase -
19	80	6.8	3161	2 T30342	protein HMWp1 - ye
20	79.5	6.8	166	2 T53242	lutreinizng hormon
21	79.5	6.8	601	2 T22025	hypothetical prote
22	78.5	6.7	1142	2 T30272	hypothetical prote
23	78.5	6.7	1469	2 B36665	silt protein 2 pre
24	78.5	6.7	2098	2 T18397	protein GTP - mal
25	78.5	6.7	3163	2 T17440	probable polyketid
26	78	6.7	834	2 T42702	hypothetical prote
27	77.5	6.6	188	2 Jc4680	vascular endotheli
28	77.5	6.6	623	2 B83399	quonoprotein alcon
29	77	6.6	623	2 S35652	transcription cont

30	77	6.6	1959	1 AGRT	agrin - rat
31	76.5	6.5	165	1 KTHUB	choirgonadotropin
32	76	6.5	647	2 JE0337	Frizzled-1 protein
33	76	6.5	865	1 D70986	probable ABC trans
34	76	6.5	1042	2 A57534	mucin 5AC (clone L
35	76	6.5	1056	2 A53767	mucin MUC5B, trach
36	76	6.5	1208	2 T27822	hypothetical prote
37	75.5	6.4	699	2 T09069	probable cAMP-resp
38	75.5	6.4	3570	2 T45025	mucin MUC5B, trach
39	75	6.4	275	2 T50578	hypothetical prote
40	75	6.4	721	2 T45495	probable transpos
41	75	6.4	1682	1 C70588	probable mte prot
42	74.5	6.4	128	2 S74085	lutropin beta chai
43	74.5	6.4	232	2 A60083	neural induction h
44	74.5	6.4	304	2 E70698	hypothetical prote
45	74	6.3	269	2 D75631	iron ABC transport

ALIGNMENTS

RESULT 1

T08710

hypothetical protein DKFZP564D206.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999

C:Accession: T08710

R:Wambolt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

Submitted to the Protein Sequence Database, March 1999

A:Reference number: Z16471

A:Accession: T08710

A:Molecule type: mRNA

A:Residues: 1-134 <WAM>

A:Cross-references: EMBL:AL050024

A:Experimental source: fetal brain; clone DKFZP564D206

C:Genetics:

A>Note: DKFZP564D206.1

Query Match 21.3%; Score 249.5; DB 2; Length 134;

Best local similarity 39.7%; Pred. No. 8e-15; Matches 52; Conservative 28; Mismatches 44; Indels 7; Gaps. 4;

QY 78 CRELHYTRFLTDGPCRSANFVELVCGGCGPARLLPNAIG--RVKWM-RPNQDFPCT 133

DB 3 CRELSTKRYSPDQCTSIPLKELVCGECLLPVLPNIGGYGTKYWSRSSQEMRCV 62

QY 134 PDPRRAQROVLCPGGAAPRSRKVRVLVASCCKRLTRFHNOSELKDFGPETAPQKGRP 193

DB 63 NDKRTRIOLQCDQGST-RTIKITVYVACKCKRTTRQHNSSHNFEKSPAKPVQHHRE 121

QY 194 RFGARAKANO 204

DB 122 R-KRAKSSK 130

RESULT 2

PN0141

lutropin beta chain - sperm whale

N:Alternate names: lutreinizng hormone beta chain

C:Species: Phseter catodon (sperm whale)

C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999

C:Accession: PN0141

R:Panikov, Y.A.; Karasev, V.S.

Biochimica 49, 1004-1018, 1984.

A:Title: Lutreinizng hormone of the sperm whale: amino acid sequence of reduced and c

A:Reference number: PN0141; MOID:84281133

A:Accession: PN0141

A:Molecule type: protein

A:Residues: 1-118 <PAN>

A>Note: article in Russian with English abstract

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone

F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.9%; Score 93; DB 2; Length 118;
Best Local Similarity 30.1%; Pred. No. 0.3;
Matches 37; Conservative 13; Mismatches 55; Indels 18; Gaps 6;
QY 78 CRELHTRFLTPGCRSAKPVELVCSGQC-GPARLLPNAIGRVKMRPNPGDFRCIPDR 136
DB 9 CRPINMTLAQVZACPVCTTFTTSTICAGYCPSMNVLPAL-----PPVPZPVCTYRQ 61
QY 137 YRAQRYOLL-CPGGAAPR-SRKRVLVASCKCKRLTRFHNSGLKDFGPTARPKR-GRKP 193
DB 62 LRFASIRLPCGPPGVPMVSPVIALSCHGCPRLSS-----SBCGGRAPLACBRSP 114
QY 194 RFG 196
DB 115 RFG 117

RESULT 3

373
Luteinizing hormone beta subunit - turkey
C:Species: Meleagris gallopavo (common turkey)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: U51373
R:Yon, S.; Foster, L.R.; Silsby, J.L.; el Halawani, M.E.; Foster, D.N.
J. Mol. Endocrinol. 14, 117-129, 1995
A:Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonadotropin-releasing hormone
A:Reference number: U51373; MUID:95290073
A:Accession: U51373
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-159 <Y00>
A:Cross-references: GB:L35519; NID:9530952; PIDN:AAA74125.1; PID:9530953
C:Genetics:
A:Gene: LH-beta
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 7.9%; Score 92.5; DB 2; Length 159;
Best Local Similarity 26.6%; Pred. No. 0.45;
Matches 33; Conservative 12; Mismatches 58; Indels 21; Gaps 5;

QY 60 GGRPHHPRDAKOVSEYSCRELHYTRFLTDGCRSAKPVELVCSGQC-GPARLLPNAIGR 119
DB 43 GGRPP-----CRPINVTVAVERKDCPCQCAVTTTACGGYCTR-----EPVYR 85
QY 120 VKWMPNGDFRCIPDRYRAQRYOLL-CPGGAAPRSRKRVLVASCKCKRLTRFHNSGLK 178
DB 86 SPLGRP-POSCTYGALRYERMALMGCPISDPRV-LLPVALSRCARCIATSDCTVQ 142
QY 179 DFGP 182
DB 143 GUGP 146

RESULT 4

PN0139
Luteolin beta chain - minke whale
N:Alternate names: luteinizing hormone beta chain
C:Species: Balaeonoptera acutorostrata (minke whale, lesser torquial)
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 08-Dec-1995
C:Accession: PN0139
R:Karasev, V.S.; Pankov, Y.A.
Biol. Zhurn. 50, 1972-1986, 1985
A:Title: Amino acid sequence of reduced and carboxymethylated alpha- and beta-subunits of luteinizing hormone
A:Reference number: PN0139
A:Accession: PN0139
A:Molecule type: protein
A:Residues: 1-118 <KAR>
A:Note: article in Russian with English abstract

C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone
F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.8%; Score 91; DB 2; Length 118;
Best Local Similarity 29.3%; Pred. No. 0.45;
Matches 36; Conservative 12; Mismatches 57; Indels 18; Gaps 6;
QY 78 CRELHTRFLTPGCRSAKPVELVCSGQC-GPARLLPNAIGRVKMRPNPGDFRCIPDR 136
DB 9 CRPINMTLAQVZACPVCTTFTTSTICAGYCPSMNVLPAL-----PPVPZPVCTYRQ 61
QY 137 YRAQRYOLL-CPGGAAPR-SRKRVLVASCKCKRLTRFHNSGLKDFGPTARPKR-GRKP 193
DB 62 LRFASIRLPCGPPGVPMVSPVIALSCHGCPRLSS-----SBCGGRAPLACBRSP 114
QY 194 RFG 196
DB 115 RFG 117

RESULT 5

A61091
Luteolin beta chain precursor - chicken
N:Alternate names: luteinizing hormone beta chain
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995
C:Accession: A61091
R:Moce, T.; Ando, H.; Ueda, T.; Kubokawa, K.; Higashinakagawa, T.; Ishii, S.
J. Mol. Endocrinol. 3, 129-137, 1989
A:Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for luteinizing hormone beta subunit
A:Reference number: A61091; MUID:89374710
A:Accession: A61091
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-158 <NOC>
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone; pituitary
F:1-39/Domain: signal sequence #status predicted <SIG>
F:46-73,62-96,65-127,77-149,111-139,129-133/Disulfide bonds: #status predicted
F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.6%; Score 89.5; DB 2; Length 158;
Best Local Similarity 23.7%; Pred. No. 0.81;
Matches 32; Conservative 11; Mismatches 49; Indels 43; Gaps 5;

QY 60 GGRPHHPRDAKOVSEYSCRELHYTRFLTDGCRSAKPVELVCSGQC-----G 108
DB 43 GGRPP-----CRPINVTVAVERKDCPCQCAVTTTACGGYCTR-----EPVYR 89
QY 109 PARLLPNAIGRVKMRPNPGDFRCIPDRYRAQRYOLL-CPGGAAPRSRKRVLVASCKCKR 167
DB 90 PP-----POSCTYGALRYERMALMGCPISDPRV-LLPVALSRCARCIATSDCTVQ 131
QY 168 LTRFHNSGLKDFGP 182
DB 132 CPMAISDCTVQGLGP 146

RESULT 6

S46947
Ribosomal protein L2 - evening primrose mitochondrion
C:Species: Mitochondrion Oenothera villaricae (evening primrose)
C>Date: 27-Jan-1995 #sequence_revision 24-Feb-1995 #text_change 13-Aug-1999
C:Accession: S46947
R:Fritz, I.; Schuster, W.
Submitted to the EMBL data library, July 1994
A:Reference number: S46947
A:Accession: S46947
A:Molecule type: DNA

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:13 ; Search time 31.13 Seconds
(without alignments)
232.185 Million cell updates/sec

Title: US-09-668-021-12
1171

Perfect score: 1 MOPSIAFCLICLVHAAFC.....KPPRGAKAKAOELENNY 211

Sequence: 1 MOPSIAFCLICLVHAAFC.....KPPRGAKAKAOELENNY 211

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	93	7.9	118 1 LSHB_PHYCA	P25330 physeter ca
2	92.5	7.9	159 1 LSHB_MELEGA	P45646 meleagris g
3	91	7.8	118 1 LSHB_BALAC	P33088 balaenopter
4	86.5	7.4	141 1 LSHB_TRIVU	O46482 trichosurus
5	82	7.0	301 1 U149_HSV11	P10233 herpes simp
6	81.5	7.0	2182 1 CAB1_RAT	O88480 rattus norv
7	81	6.9	470 1 KLF4_HUMAN	O43474 homo sapien
8	80.5	6.9	348 1 GCR_SYNY3	P74034 synchocyst
9	80.5	6.9	975 1 CDP_CANTF	P39881 canis fam11
10	80	6.8	141 1 LSHB_CERSI	O77833 ceratotheri
11	79.5	6.8	888 1 KLTR_MOUSE	P08923 mus musculu
12	79.5	6.7	166 1 LSHB_COTJA	P05657 coturnix co
13	78.5	6.7	625 1 DUS8_HUMAN	O13202 homo sapien
14	78	6.7	1942 1 Y054_HUMAN	P42694 homo sapien
15	77.5	6.6	138 1 LSHB_MACRU	O46483 macropus ru
16	77.5	6.6	188 1 VEGB_MOUSE	P49766 mus musculu
17	77.5	6.6	623 1 EXAA_PSEAE	Q92437 pseudomona
18	77	6.6	1959 1 AGR1_RAT	P25304 rattus norv
19	76.5	6.5	165 1 CGBH_HUMAN	P01233 homo sapien
20	76.5	6.5	170 1 PLGF_HUMAN	P49764 homo sapien
21	76	6.5	143 1 LSHB_FELCA	O77805 felis silve
22	76	6.5	644 1 VP4_BRV11	P33428 bluetongue
23	76	6.5	644 1 VP4_BRV13	P33429 bluetongue
24	76	6.5	1056 1 MUC5_HUMAN	P38088 homo sapien
25	74.5	6.4	128 1 LSHB_STRCA	P80664 struthio ca
26	74.5	6.4	232 1 HXB9_XENLA	P11272 xenopus lae
27	74	6.3	180 1 DAN_HUMAN	P41271 homo sapien
28	74	6.3	379 1 CYR6_MOUSE	P18406 mus musculu
29	74	6.3	381 1 CYR6_HUMAN	O00622 homo sapien
30	74	6.3	509 1 CRO3_BRANA	P33523 brassica na
31	73.5	6.3	1480 1 SLIT_DROME	P24014 drosophila
32	73.5	6.3	174 1 BARI_CHITE	P02849 chironomus
33	73.5	6.3	474 1 KLF4_MOUSE	O60793 mus musculu

34	73.5	6.3	654 1 VP4_BRV10	P07132 bluetongue
35	73.5	6.3	969 1 SYL_MYCTU	P71698 mycobacteri
36	73.5	6.3	1172 1 LMB3_HUMAN	O13751 homo sapien
37	73	6.2	178 1 DAN_RAT	O06880 rattus norv
38	73	6.2	992 1 ERN6_EBV	P03204 Epstein-Bar
39	73	6.2	1186 1 CEN4_BACTS	Q45710 bacillus th
40	73	6.2	1955 1 AGR1_CHICK	P31696 gallus galli
41	73	6.2	2481 1 UN52_CAEL	O06561 caenorhabdi
42	72.5	6.2	628 1 MSIN_HUMAN	O13421 homo sapien
43	72	6.1	401 1 HMEL_MOUSE	P09065 mus musculu
44	72	6.1	532 1 ZN20_HUMAN	P17024 homo sapien
45	72	6.1	890 1 AT8S_HUMAN	G9up79 homo sapien

ALIGNMENTS

RESULT	ID	STANDARD	PRT	118 AA.
1	LSHB_PHYCA			
AC	P25330:			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	LUTROPIN BETA CHAIN (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA)			
DE	(LSH-B) (LH-B).			
GN	LHB.			
OS	Physeter catodon (Sperm whale) (Physeter macrocephalus).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;			
OC	Physeteridae; Physeter.			
OX	NCBI_TaxId=9755;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=87032654; PubMed=37711098;			
RA	Pankov Y.A., Karasyov V.S.;			
RT	"Primary structure of sperm whale luteinizing hormone.";			
RT	Int. J. Pept. Protein Res. 28:124-129(1986).			
RN	[2]			
RP	SEQUENCE.			
RX	MEDLINE=84281133; PubMed=6466737;			
RA	Pankov Y.A., Karasev V.S.;			
RT	"Luteinizing hormone of the sperm whale. Amino acid sequences of reduced and carboxymethylated beta-subunits.";			
RT	Biochimica 49:1004-1018(1984).			
CC	-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.			
CC	-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.			
CC	-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.			
CC	PIR: PNO141; PNO141.			
DR	HSSP: P01233; 1HRP.			
DR	InterPro: IPR000359; -			
DR	InterPro: IPR001545; -			
DR	PIfam: PF00007; Cys_knot; 1.			
DR	PRINTS: PR00438; GFCYSKNOT.			
DR	PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.			
DR	PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.			
KW	Hormone; Glycoprotein.			
FT	DISULFID 9	57		BY SIMILARITY.
FT	DISULFID 23	72		BY SIMILARITY.
FT	DISULFID 26	110		BY SIMILARITY.
FT	DISULFID 34	88		BY SIMILARITY.
FT	DISULFID 38	90		BY SIMILARITY.
FT	DISULFID 93	100		BY SIMILARITY.
FT	CARBOHYD 13	13		N-LINKED (GLCNAC. . .).
SQ	SEQUENCE 118 AA; 12412 MW; 8117756382F15E7 CRC64;			

Query Match 7.9%; Score 93; DB 1; Length 118;

Best Local Similarity 30.1%; Pred. No. 0.089;
Matches 37; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 78 CRELHYTRFLDGPGRSAKPYTELVCSGGOC-GPARLLPNAIGRVMKPNNGDFRCIDPR 136
DB 9 CRINMTLANQACACVCTFTTSTGAGCPSMVRVLPAL-----PPVZPCTYTRQ 61
QY 137 YRAQVOLL-CPGGAAPR-SRKVRLVASCCKRLTRFNHOSLKDGPETAPQR-GRRP 193
DB 62 LRFASIRLPGCPGVNPMVSPVALSCHGCPRLSS-----SDCGPGRAPLACNRRSP 114
QY 194 RFG 196
DB 115 RFG 117

RESULT 2
LSHB_MEIGA
ID LSHB_MEIGA STANDARD; PRT; 159 AA.
AC P45646;

01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).

OS Melagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
OX NCBI_Taxid=9103;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Plutary;
RX MEDLINE=95290073; PubMed=7772235;
RA You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;
RT Sequence analysis of the turkey LH beta subunit and its regulation
RT by gonadotrophin-releasing hormone and prolactin in cultured
RT pituitary cells.";

RL J. Mol. Endocrinol. 14:117-129(1995).

-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.

-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: L35519; AAA74125.1; ALT_INIT.

DR HSSP: P01233; IHRP.

DR InterPro: IPR000359; -

DR InterPro: IPR001545; -

DR Pfam: PF00007; Cys_knot; 1.

DR PRINTS: PR00438; GFCYSKNOT.

DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.

DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.

KW Hormone; Signal; Glycoprotein.

FT SIGNAL 1 39

FT CHAIN 40 159

FT DISULFID 48 96

FT DISULFID 62 111

FT DISULFID 65 149

FT DISULFID 73 127

FT DISULFID 77 129

FT DISULFID 132 139

FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 159 AA; 16285 MW; 52850C8C879653C6 CRC64;

Query Match 7.8%; Score 92.5; DB 1; Length 159;

Best Local Similarity 26.6%; Pred. No. 0.13;
Matches 33; Conservative 12; Mismatches 58; Indels 21; Gaps 5;

QY 60 GGRPHHPYDAKVSEYSCRELHYTRFLDGPGRSAKPYTELVCSGGOC-GPARLLPNAIGR 119
DB 43 GGRPP-----CRPINVTVAVERKDECPQCMATVTAACGGYCTR-----EPYR 85

QY 120 VKMWRPNPDPFCIDPRYRAQVOLL-CPGGAAPRSRKRLVASCCKRLTRFNHOSLKD 178
DB 86 SPLGRP--POSSCTYGALRYERWALMGCPISDPRV-LLPVALSCRCARCPATSDCTVQ 142

QY 179 DFGP 182
DB 143 GLGP 146

RESULT 3
LSHB_BALAC
ID LSHB_BALAC STANDARD; PRT; 118 AA.
AC P33088;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DE LUTROPIN BETA CHAIN (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA)
(LSH-B) (LH-B).

GN Balaeoptera acutirostrata (Mink whale) (lesser rorqual).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;

OC Balaeopteridae; Balaeoptera.

OX NCBI_Taxid=9767;

RN [1]

RP SEQUENCE.

RA Karasev V.S., Pankov Y.A.;

RT "Amino acid sequence of reduced and carboxymethylated alpha- and beta-
RT subunits of the little picked whale luteinizing hormone.";

RL Biochimica 50:1972-1986(1985).

-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.

-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
FAMILY.

DR PIR: P0135; P0139.

DR HSSP: P01233; IHRP.

DR InterPro: IPR000359; -

DR InterPro: IPR001545; -

DR Pfam: PF00007; Cys_knot; 1.

DR PRINTS: PR00438; GFCYSKNOT.

DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.

DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; FALSE_NEG.

KW Hormone; Glycoprotein.

FT DISULFID 9 57

FT DISULFID 23 72

FT DISULFID 26 110

FT DISULFID 34 88

FT DISULFID 38 90

FT DISULFID 93 100

FT CARBOHYD 13 13

FT SEQUENCE 118 AA; 12414 MW; 039F229EFC480F5D CRC64;

Query Match 7.8%; Score 91; DB 1; Length 118;
Best Local Similarity 29.3%; Pred. No. 0.13;
Matches 36; Conservative 12; Mismatches 57; Indels 18; Gaps 6;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:00:35 ; Search time 89.34 Seconds
(without alignments)
312.473 Million cell updates/sec

Title: US-09-668-021-12
Perfect score: 1171
Sequence: 1 MQRSLAPCLICILYHAFFCA.....KRPFGARGAQANQAELENNY 211

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

rchd: 425026 seqs, 13205027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-unclassified:*
13: sp-vertebrate:*
14: sp-virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249.5	21.3	134	4 Q9Y3U3	Q9Y3U3 homo sapien
2	90.5	7.7	184	3 Q73755	Q73755 gallus gall
3	90	7.7	712	4 Q9Y557	Q9Y557 homo sapien
4	90	7.7	720	4 Q9NY23	Q9NY23 homo sapien
5	90	7.7	988	6 Q97867	Q97867 sus scrofa
6	89	7.6	215	8 Q9MFB6	Q9MFB6 beta vulgar
7	89	7.6	261	2 Q9F1E4	Q9F1E4 acetobacter
8	89	7.6	182	13 Q13091	Q13091 pleurodeles
9	88.5	7.6	184	4 Q60565	Q60565 homo sapien
10	88.5	7.6	184	11 Q35793	Q35793 ratus norv
11	88	7.5	281	2 Q9EMD9	Q9EMD9 streptomyc
12	88	7.5	332	10 Q40600	Q40600 oenothera b
13	88	7.5	272	13 Q9PWB0	Q9PWB0 gallus gall
14	87	7.4	482	14 Q71654	Q71654 human immun
15	86.5	7.4	1114	11 Q9JKW7	Q9JKW7 mus musculu
16	86.5	7.4	184	11 Q70326	Q70326 mus musculu
17	86	7.3	168	4 Q9H772	Q9H772 homo sapien
18	85	7.3	272	13 Q9PUK2	Q9PUK2 gallus gall
19	85	7.3	272	13 Q9PUK2	Q9PUK2 gallus gall

20	84.5	7.2	791	11 Q70368	Q70368 ratus norv
21	84.5	7.2	1223	11 Q9QW33	Q9QW33 ratus sp.
22	84	7.2	168	11 Q88273	Q88273 mus musculu
23	83	7.1	1035	5 Q9NEG1	Q9NEG1 drosophila
24	82.5	7.0	334	10 Q49658	Q49658 arabidopsis
25	82	7.0	1081	4 Q76065	Q76065 homo sapien
26	82	7.0	1537	11 Q9WTW6	Q9WTW6 mus musculu
27	82	7.0	1581	11 Q9R0B6	Q9R0B6 mus musculu
28	82	7.0	1637	6 Q9XSV8	Q9XSV8 bos taurus
29	81.5	7.0	349	10 P93311	P93311 arabidopsis
30	81.5	7.0	13288	6 Q18758	Q18758 sus scrofa
31	81	6.9	470	4 Q9UNP3	Q9UNP3 homo sapien
32	81	6.9	1242	4 Q9NS15	Q9NS15 homo sapien
33	81	6.9	1382	4 Q9H7K2	Q9H7K2 homo sapien
34	80.5	6.9	270	13 P70041	P70041 xenopus lae
35	80.5	6.9	1704	5 Q94446	Q94446 chironomus
36	80	6.8	267	4 Q95813	Q95813 homo sapien
37	80	6.8	866	5 P91426	P91426 caenorhabdi
38	80	6.8	1212	2 Q911C8	Q911C8 streptomyc
39	80	6.8	3161	2 Q54511	Q54511 yersinia en
40	79.5	6.8	1616	4 Q15054	Q15054 homo sapien
41	78.5	6.7	641	5 Q97355	Q97355 drosophila
42	78.5	6.7	641	5 Q9YVW6	Q9YVW6 drosophila
43	78.5	6.7	648	4 Q9UP38	Q9UP38 homo sapien
44	78.5	6.7	818	4 Q9Z0M6	Q9Z0M6 mus musculu
45	78.5	6.7	1142	5 Q26615	Q26615 strongyloce

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	134 AA.
Q9Y3U3	Q9Y3U3			
AC	Q9Y3U3:			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	HYPOTHEITICAL 15.3 KDA PROTEIN (FRAGMENT).			
GN	DKZP564D206.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RA	Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;			
RL	Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AL050024; CAB43243.1; -.			
DR	InterPro; IPR000359; -.			
DR	PROSITE; PS01225; CTCK_2; 1.			
KW	Hypothetical protein.			
FT	NON_TER 1			
SO	SEQUENCE 134 AA; 15324 MW; A0B7A8650D2EE6F1 CRC64;			

Query Match 21.3%; Score 249.5; DB 4; Length 134;
Best Local Similarity 39.7%; Pred. No. 2.3e-17;
Matches 52; Conservative 28; Mismatches 44; Indels 7; Gaps 4;

QY	78 CRELHYRFLTDGPCRSKAPYTELVCSGCGPARLPLNAIG---RKQMW-RPMGPRFRCI 133
DB	3 CRELSTFYISDGOCTSIPLKELVCAEGECLLPVLPNWTGGYGRKYWSSQSWRCV 62
QY	134 PDRRAQVOLLCPGAPRGRKRLVASCCKRLRFRHNSSELDKDPGRTAPRQGRKP 193
DB	63 NDKRTGQIQLQCDGSR-KTYKTTVTACKCKRYTRQHNHSSINFSMSAPRQVQHNHE 121
QY	194 RFGARGAQANQ 204
DB	122 R--KRASKSK 130

RESULT 2
 ID 073755 PRELIMINARY; PRT; 184 AA.
 AC 073755;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GREMLIN.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 RN NCBL_TaxID=9031;
 RP [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=9835381; PubMed=9660951;
 RA Hsu D.R., Economides A.N., Wang X., Eimon P.M., Harland R.M.;
 RT "The Xenopus dorsaling factor Gremlin identifies a novel family of
 secreted proteins that antagonize BMP activities.";
 RL Mol. Cell. 1:673-683(1998).
 DB EMBL; AF045799; AAC41280.1; -.
 DR InterPro; IPR000359; -.
 SM SMART; SM00041; CT; 1.
 SQ SEQUENCE 184 AA; 21166 MW; 3510B4E8F6DD5EA CRC64;

Query Match 7.7%; Score 90.5; DB 13; Length 184;
 Best Local Similarity 22.9%; Pred. No. 0.19;
 Matches 35; Conservative 23; Mismatches 60; Indels 35; Gaps 8;

QY 41 GEXPEP-----PENNOTMRAENGRRPHH-----PYDAKVSEYSGRELHYT--RFLTDG 90
 DB 33 GAIPEPDKQPNDSSEMQQSGSHRERCKGTMPAEVLESQELHITERKYLKRD 92
 QY 91 PCR-----SAKPVELVCSGGCGPARLLPNAIGRWKWRPNDPFR-----CI 133
 DB 93 WCKTOPLKQTIHEEGCNSTIIRFCYGCN--SFYIPRHV-----RKEGSEFOSCFCK 145
 QY 134 PDREARQRYQLCPGAPAPRSRK--VRLVASCCK 165
 DB 146 PKFTTMYTLNCPLOPRKKKRITRYKRCRC 178

RESULT 3
 ID 09Y557 PRELIMINARY; PRT; 712 AA.
 AC 09Y557;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DT DJ1163J1.3 (NOVEL PROTEIN SIMILAR TO MOUSE B99) (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN NCBL_TaxID=9606;
 RP [1]
 SEQUENCE FROM N.A.
 RX Lloyd D.;
 RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL031588; CAB38415.1; -.
 FT NON_TER 1
 SQ SEQUENCE 712 AA; 75685 MW; ECB2D1333EF72F32 CRC64;

Query Match 7.7%; Score 90; DB 4; Length 712;
 Best Local Similarity 22.7%; Pred. No. 0.83;
 Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;
 QY 91: PCSAKPVELVCSGGCGPARLLP-----NAIGRWKWRPNDPFRICIDRYRAQRY----- 142
 DB 333 PANSSRPLSNISKSGRMGPALRLPALPAGPVGASSW-----QAKRVDSVSE 377

QY 143 ----QLCPGGAAP-----RSKRYRLVASC---CKRLT 169
 DB 378 LAABQLTAPPSASPTQPTPEGGGQWLNSSCAWSESSQLNKRIRRDSCINRKYMP 437
 QY 170 RFHNSLKDPE-----GPEARPOKGRKRP 195
 DB 438 TPTMQFKTPKFSIGDSPDSTPLKRAQRP 467

RESULT 4
 ID 09NY23 PRELIMINARY; PRT; 720 AA.
 AC 09NY23;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE B99.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN NCBL_TaxID=9606;
 RP [1]
 SEQUENCE FROM N.A.
 RA Monte M., Collavin L., Lazarevic D., Utrera R., Schneider C.;
 RT "Cloning, chromosome mapping and functional characterization of a
 human homolog gene of murine B99";
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF223408; AAF31459.1; -.
 SQ SEQUENCE 720 AA; 76628 MW; D4E852CEABE85BDF CRC64;

Query Match 7.7%; Score 90; DB 4; Length 720;
 Best Local Similarity 22.7%; Pred. No. 0.84;
 Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 91 PCSAKPVELVCSGGCGPARLLP-----NAIGRWKWRPNDPFRICIDRYRAQRY----- 142
 DB 341 PANSSRPLSNISKSGRMGPALRLPALPAGPVGASSW-----QAKRVDSVSE 385
 QY 143 ----QLCPGGAAP-----RSKRYRLVASC---CKRLT 169
 DB 386 LAABQLTAPPSASPTQPTPEGGGQWLNSSCAWSESSQLNKRIRRDSCINRKYMP 445
 QY 170 RFHNSLKDPE-----GPEARPOKGRKRP 195
 DB 446 TPTMQFKTPKFSIGDSPDSTPLKRAQRP 475

RESULT 5
 ID 097867 PRELIMINARY; PRT; 988 AA.
 AC 097867;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GASTRIC MUCIN (FRAGMENT).
 GN MUC5AC.
 OS Sus scrofa domestica (domestic pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 RN NCBL_TaxID=9825;
 RP [1]
 SEQUENCE OF 35-109 FROM N.A.
 RC TISSUE=GASTRIC EPITHELIUM;
 RX MEDLINE=95275264; PubMed=7755593;
 RA Turner B.S., Bhaskar K.R., Hadzopolou-Ciadaras M., Specian R.D.,
 RT Lamort J.T.;
 RT "Isolation and characterization of cDNA clones encoding pig gastric
 mucin";
 RL Biochem. J. 308:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:42 ; Search time 80.28 Seconds
(without alignments)
159.338 Million cell updates/sec

Title: US-09-668-021-12

Perfect score: 1171
Sequence: 1 MGPLAPLCICLLVHAAFCA.....KPRGARGAKANQALENAY 211

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1171	100.0	211	21	AAV96432 Murine TGF-beta b1
2	1094	93.4	213	21	AAV96433 Rat TGF-beta bindi
3	1052	89.8	213	21	AAV96431 Vervet TGF-beta b1
4	1029	87.9	213	21	AAV96430 Human DAN/Cerberus
5	1029	87.9	213	21	AAV96429 Human TGF-beta bin
6	1029	87.9	213	21	AAV97389 Human secreted pro
7	1021	87.2	213	21	AAV96430 Human TGF-beta bin
8	1020	87.1	213	21	AAV96436 Human TGF-beta bin
9	934	79.8	176	21	AAV96434 Bovine DAN/Cerberus
10	857.5	73.2	167	21	AAV96434 Bovine TGF-beta b1
11	340.5	29.1	206	21	AAV75981 Murine skin cell p

12	340.5	29.1	206	21	AAV76031 Murine skin cell p
13	340.5	29.1	206	22	AAV55920 Skin cell protein,
14	340.5	29.1	206	22	AAV55970 Skin cell protein,
15	333	28.4	206	18	AAW05408 Human small CCN-11
16	333	28.4	206	19	AAW58704 Human small CCN-11
17	333	28.4	206	20	AAW95711 Homo sapiens fetal
18	333	28.4	206	21	AAV10233 Human adult retina
19	215	18.4	50	20	AAV12009 Human 5' EST secre
20	114	9.7	116	18	AAW27654 Secreted protein A
21	114	9.7	116	18	AAW44090 Human secreted pro
22	98.5	8.4	102	20	AAV03225 Amino acid sequenc
23	98.5	8.4	102	21	AAV10277 Human fetal kidney
24	90.5	7.7	184	19	AAW29727 DAN and b57 protei
25	90	7.7	712	21	AAV43147 Human ORFX ORF2911
26	88.5	7.6	182	19	AAW29726 DAN and b57 protei
27	88.5	7.6	184	19	AAW6293 Human b57 protein
28	88.5	7.6	184	20	AAV42173 Mouse DRM protein
29	88.5	7.6	184	21	AAV95961 Human gremlin. Ho
30	88	7.5	184	20	Rat DRM protein se
31	88	7.5	429	20	EGFP/DRM fusion pr
32	87	7.4	391	20	AAV42178 Human DRM protein
33	86	7.3	184	20	AAV42172 Human DRM protein
34	86	7.3	212	21	AAV43169 Human ORFX ORF2933
35	86	7.3	301	21	AAV83261 HSV-1 V22 cellular
36	85	7.3	168	21	AAV84014 Amino acid sequenc
37	84	7.2	168	21	AAV51132 Murine cerebral ne
38	83.5	7.1	165	20	AAW99515 Glycoprotein hormo
39	83	7.1	386	18	AAW18664 Fragmented human N
40	82.5	7.0	145	12	AAV15109 hCG/ELH chimera, E
41	82.5	7.0	304	21	AAV20509 Arabidopsis thalia
42	82.5	7.0	304	21	AAV35061 Arabidopsis thalia
43	82.5	7.0	334	21	AAV20508 Arabidopsis thalia
44	82.5	7.0	334	21	AAV35060 Arabidopsis thalia
45	82	7.0	258	22	AAV88367 Human membrane or

ALIGNMENTS

RESULT 1	AAV96432 standard. Protein; 211 AA.
ID	AAV96432
AC	AAV96432
XX	12-SEP-2000 (first entry)
DT	12-SEP-2000 (first entry)
XX	Murine TGF-beta binding protein (BEER).
DE	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW	BEER, gene therapy; antisense therapy; fracture; bone mineralization.
OS	Mus musculus.
XX	WO200032773-A1.
PN	08-JUN-2000.
PD	24-NOV-1999; 99WO-US27990.
PF	27-NOV-1998; 98US-0110283.
PR	(DARW-) DARWIN DISCOVERY LTD.
XX	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
PI	Van Ness J, Winkler DG;
XX	WPI; 2000-412321/35.
DR	N-PSDB; AAA29058.
XX	Nucleic acids (I) encoding a transforming growth factor beta binding
PT	protein, useful for identifying agents for treating osteopenia,
PT	osteoporosis and fractures

XX Claim 5; Page 124; 162pp; English.
PS
XX
CC This shows a murine transforming growth factor-beta (TGF-beta)
CC binding protein designated mBER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patients own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.
SQ Sequence 211 AA;
XX

Query Match 100.0%; Score 1171; DB 21; Length 211;
Best Local Similarity 100.0%; Pred. No. 5.5e-100;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOPSLAPCLICILVHAFCAGVGOGQAFRNDATEVTPGLGEYPPPPNNQTMNRAENG 60
Db 1 mgpslapclilcllvhafeavegvgqatrndatevlp91geyppeppennqumraeng 60
QY 61 GRPHHPHYAKDVSEYSECRELHTRFLTDGPCRSAPVTELVCSGCCGPARLLPNAIGRV 120
Db 61 grphhpyakadvseyscrelhytrfltdgpcrsakpvtelvcsgcgparllpnaigrv 120
QY 121 KMKRPNPDPFRCLPDYRRAQROVOLLCPGGAAPRSKRVRLVASCCKRLTRFNHOSLEKPF 180
Db 121 kwmrpnppdfrclpdyrraqrvollcp9gaaprskrvrlvasckckrltrfnhngselkpf 180
QY 181 GPETARPOGKRPRPGARGAKANOAELENAY 211
Db 181 gpetarpgkyrkprpgargakanaeelenay 211

RESULT 2
ID AAY96433
AC AAY96433;
XX
DT 12-SEP-2000 (first entry)
XX
Rat TGF-beta binding protein (BEER).
XX
KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX
OS Rattus norvegicus.
XX
PN MO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
XX (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
XX Van Ness J, Winkler DG;
XX
DR WPI; 2000-412321/35.
DR N-PSDB; AAA29059.

XX Nucleic acids (1) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
PS
XX Claim 6; Page 125-126; 162pp; English.
XX
CC This shows a rat transforming growth factor-beta (TGF-beta) binding
CC protein designated rBER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patients own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.
SQ Sequence 213 AA;
XX

Query Match 93.4%; Score 1094; DB 21; Length 213;
Best Local Similarity 93.9%; Pred. No. 6.5e-93;
Matches 200; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MOPSLAPCLICILVHAFCAGVGOGQAFRNDATEVTPGLGEYPPPP--ENNQTMNRAE 58
Db 1 mgslapclilcllvhafeavegvgqafkndatevlp91reyepqgelennqumrae 60
QY 59 NCGRPNHPHYAKDVSEYSECRELHTRFLTDGPCRSAPVTELVCSGCCGPARLLPNAIG 118
Db 61 nggrpnhpyakadvseyscrelhytrfltdgpcrsakpvtelvcsgcgparllpnaig 120
QY 119 RKKMWRPNPDPFRCLPDYRRAQROVOLLCPGGAAPRSKRVRLVASCCKRLTRFNHOSLEK 178
Db 121 rkkwmrpnppdfrclpdyrraqrvollcp9gaaprskrvrlvasckckrltrfnhngselk 180
QY 179 DFGPETARPOGKRPRPGARGAKANOAELENAY 211
Db 181 dfgpetarpgkyrkprpgargakanaeelenay 213

RESULT 3
ID AAY96431
AC AAY96431;
XX
DT 12-SEP-2000 (first entry)
XX
Verret TGF-beta binding protein (BEER).
XX
KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX
OS Cercopithecus pygerythrus.
XX
PN MO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
XX (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:47:17 ; Search time 47.86 Seconds
(without alignments)
99.210 Million cell updates/sec

Title: US-09-668-021-12

Sequence: 1 M0PSLAPCLICLLVHAFACA.....KPRPGARAKANQAELENAY 211

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	340.5	29.1	206	4	US-09-188-930-159 Sequence 159, App
2	340.5	29.1	206	4	US-09-188-930-286 Sequence 286, App
3	333	28.4	206	1	US-08-468-847B-2 Sequence 2, Appl
4	333	28.4	206	1	US-08-468-847B-20 Sequence 20, Appl
5	82	7.0	301	3	US-08-303-861-21 Sequence 21, Appl
6	82	7.0	301	4	US-09-011-073A-1 Sequence 1, Appl
7	82	7.0	301	4	US-09-230-421-2 Sequence 2, Appl
8	81	6.9	405	5	PCT-US93-11404-2 Sequence 2, Appl
9	81	6.9	470	2	US-08-946-241B-2 Sequence 2, Appl
10	81	6.9	470	3	US-09-309-053-2 Sequence 2, Appl
11	81	6.9	479	2	US-08-946-241B-9 Sequence 9, Appl
12	81	6.9	479	3	US-09-309-053-9 Sequence 9, Appl
13	80.5	6.9	270	4	US-08-878-474-1 Sequence 9, Appl
14	79.5	6.8	735	3	US-09-191-647-9 Sequence 9, Appl
15	79.5	6.8	735	4	US-09-540-245A-9 Sequence 9, Appl
16	79.5	6.8	735	4	US-09-540-153-9 Sequence 9, Appl
17	79	6.7	1213	4	US-09-413-814-79 Sequence 79, Appl
18	77.5	6.6	188	1	US-08-468-427A-5 Sequence 5, Appl
19	77.5	6.6	188	2	US-08-609-443B-5 Sequence 5, Appl
20	77.5	6.6	188	2	US-08-569-063C-5 Sequence 5, Appl
21	77.5	6.6	195	2	US-08-469-427A-7 Sequence 7, Appl
22	77.5	6.6	195	2	US-08-609-443B-7 Sequence 7, Appl
23	77.5	6.6	195	2	US-08-569-063C-7 Sequence 7, Appl
24	77	6.6	55	1	US-08-468-427A-3 Sequence 3, Appl
25	77	6.6	55	2	US-08-609-443B-3 Sequence 3, Appl
26	77	6.6	55	2	US-08-569-063C-3 Sequence 3, Appl
27	77	6.6	1940	2	US-08-644-271-30 Sequence 30, Appl

28	76.5	6.5	145	1	US-08-298-189B-1 Sequence 1, Appl
29	76.5	6.5	145	1	US-08-475-213-10 Sequence 10, Appl
30	76.5	6.5	145	2	US-08-395-238-2 Sequence 2, Appl
31	76.5	6.5	145	4	US-09-142-320-12 Sequence 12, Appl
32	76.5	6.5	145	4	US-09-142-320-13 Sequence 13, Appl
33	76.5	6.5	145	4	US-09-142-320-14 Sequence 14, Appl
34	76.5	6.5	145	4	US-09-142-320-15 Sequence 15, Appl
35	76.5	6.5	145	4	US-08-918-288-68 Sequence 68, Appl
36	76.5	6.5	145	4	US-08-282-357-68 Sequence 68, Appl
37	76.5	6.5	265	4	US-08-918-288-39 Sequence 39, Appl
38	76.5	6.5	265	4	US-08-918-288-39 Sequence 39, Appl
39	76.5	6.5	265	4	US-09-282-357-39 Sequence 39, Appl
40	76.5	6.5	265	4	US-09-282-357-39 Sequence 39, Appl
41	76.5	6.5	307	4	US-08-804-166-4 Sequence 4, Appl
42	76.5	6.5	307	4	US-08-910-991-4 Sequence 4, Appl
43	76.5	6.5	336	4	US-08-804-166-8 Sequence 8, Appl
44	76.5	6.5	336	4	US-08-910-991-8 Sequence 8, Appl
45	75.5	6.4	165	2	US-08-709-924-2 Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-188-930-159
; Sequence 159, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.101CI
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: mouse
; US-09-188-930-159

Query Match      29.1%  Score 340.5; DB 4; Length 206;
Best Local Similarity 37.7%  Pred. No. 8.9e-28;
Matches 80; Conservative 38; Mismatches 71; Indels 23; Gaps 9;

QY 4 SLAPCLICLLVHAFACAVEGQWQAFRDATETVTPGIGVPEPP-PENNQTNNRAENGGR 62
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 9 SLIP-LLICILMRNCL-----AFKNDATETLYSHVYKPVFAHPSNSTLNQNRNGR 58

QY 63 PPHNP---YDAKDSEVSECRRLTTRFLTDGPRRSKAPVTLYSCGQCPCALLNATG- 118
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 59 --HFSSTLIDNSNRVQCCRELSTFKYISDQCTSTSPKELVCAECCLPLVLNWTGG 116

QY 119 --RYVGM--RPGCPDFRCIPDRYRAQRYOLLCPGAAPRSKRYLVAASCCKRLTRFNQDS 175
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 117 GYGTWYWRSSQEMRCVNDKTRFQRLQCCDGSF--RTYKTIIVYACCKRYTQHNHDS 175

QY 176 ELKDGPEPTARPKGRKPRGARGAKANQAE 207
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 176 SHNFESVSPAKPAQHHRER--KRASKSKSHSL 205

RESULT 2
US-09-188-930-286
; Sequence 286, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
```

APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Orrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011cl
CURRENT APPLICATION NUMBER: US/09/188.930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 286
LENGTH: 206
TYPE: PRT
ORGANISM: Mouse
US-09-188-930-286

Query Match

29.1%; Score 340.5; DB 4; Length 206;

Best Local Similarity 37.7%; Pred. No. 8.9e-28; Mismatches 71; Indels 23; Gaps 9;

Matches 80; Conservative 38; Indels 23; Gaps 9;

4 SLAPCLICLVHAACVAGCGOAFRNDATFVIGLGEYPEPP-PENNOTMRAENGGR 62

9 SLIP-LCLIMNCL-----AFKNDATFELXSHVVKVPAHPSNSTLNQARNGR 58

63 PPHNP---YDADVSEYSCRELHYTRFLTDGPCNSAKVTELVCSGCCGPARLLPNAIG- 118

59 --HFSTGLDRMSRVQVGRSLRSTKYISDQCTISPLKELVCAGCELPVLPWIGGYG 116

119 --RVKMW-RPNGDFRCIPDRYRAORVOLLCPGAAPRSKRVLRVASCCKRLTRPHNOS 175

117 GYGTKTKWSRSSQEMRCVMDKTRTORIOLQODGST-RTYITVYVYACKCKRTTQHNES 175

176 ELKDPETAPRQKGRKPPGAKAKANOAEI 207

176 SHNESVSPAKPAOHRRER--KRASKSKHSL 205

Db

RESULT 3

US-08-468-847B-2

Sequence 2, Application US/08468847B

Patent No. 5780263

GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.

TITLE OF INVENTION: Human CCN-Like Growth Factor

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468, 847B

FILING DATE: 6 June 1995

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-2

Query Match

28.4%; Score 333; DB 1; Length 206;

Best Local Similarity 37.3%; Pred. No. 5.3e-27; Mismatches 76; Conservative 39; Indels 22; Gaps 8;

9 LICLLVHAACVAGCGOAFRNDATFVIGLGEYPEPP-PENNOTMRAENGGRPPHP 67

13 LACILMKSCLE-----AFKNDATFELXSHVVKVPAHPSNSTLNQARNGR--HFS 61

68 YDADVSEYSCRELHYTRFLTDGPCNSAKVTELVCSGCCGPARLLPNAIG---RVK 121

62 NTGLDRMTRVQVGRSLRSTKYISDQCTISPLKELVCAGCELPVLPWIGGYG 121

122 MW-RPNGDFRCIPDRYRAORVOLLCPGAAPRSKRVLRVASCCKRLTRPHNOS 180

122 YWSRSSQEMRCVMDKTRTORIOLQODGST-RTYITVYVYACKCKRTTQHNES 180

181 GPTAPRQKGRKPPGAKAKANO 204

181 SMSPAKPVQOHRRER--KRASKSK 202

Db

RESULT 4

US-08-468-847B-20

Sequence 20, Application US/08468847B

Patent No. 5780263

GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.

TITLE OF INVENTION: Human CCN-Like Growth Factor

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468, 847B

FILING DATE: 6 June 1995

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:

TELEFAX: 201-994-1700

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 206 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 8.2%; Score 95.5; DB 2; Length 159;
Best Local Similarity 27.4%; Pred. No. 0.19;
Matches 34; Conservative 11; Mismatches 58; Indels 21; Gaps 5;

QY 62 GGRPPHPRDVKVSEYSCRELHYRFTVDGRCRSKAPVTELVCSGCGPAPLIPNAIGR 121
|||
DB 43 GGRPP-----CRPINVYVAVEKDCPCQMAVTTTACGCGYCTR---EPYR 85
QY 122 VKWVRNGDFRCIPDRYRAQRVOLL-CPGGAAPRSKRVLVASCCKRLTFRHNSGLK 180
|||
DB 86 SPLGRF--FQSSCTYALAKYERMLMGCPIGSDPRV-LLPVALSCRCARCPATSDCIYQ 142
QY 181 DEGP 184
||
DB 143 GLGP 146

RESULT 3

lutropin beta chain precursor - chicken

N:Alternate names: luteinizing hormone beta chain
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995
C:Accession: A61091
R:Moce, T.; Ando, H.; Ueda, T.; Kubokawa, K.; Higashinakagawa, T.; Ishii, S.
J. Mol. Endocrinol. 3, 129-137, 1989
A:Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for the
A:Reference number: A61091; MUID:89374710
A:Accession: A61091
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-158 <NOC>
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone; pituitary
F:1-39/Domain: signal sequence #status predicted <SIG>
F:48-73,62-96,65-177,77-149,111-139,129-132/Disulfide bonds: #status predicted
F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.9%; Score 92.5; DB 2; Length 158;
Best Local Similarity 24.4%; Pred. No. 0.36;
Matches 33; Conservative 10; Mismatches 49; Indels 43; Gaps 5;

QY 62 GGRPPHPRDVKVSEYSCRELHYRFTVDGRCRSKAPVTELVCSGCG-----G 110
|||
DB 43 GGRPP-----CRPINVYVAVEKDCPCQMAVTTTACGCGYCTRREPVYRSPUG 89
QY 111 PARLLPNAIGRVKWRPNRNGDFRCIPDRYRAQRVOLL-CPGGAAPRSKRVLVASCCKR 169
|||
DB 90 PP-----FQSACTYALAKYERMLMGCPIGSDPRV-LLPVALSCRCAR 131
QY 170 LTRFHNSGLKDFGP 184
||
DB 132 CPMATSDCTVOGLGP 146

RESULT 4

lutropin beta chain - sperm whale

N:Alternate names: luteinizing hormone beta chain
C:Species: Physter catodon (sperm whale)
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999
C:Accession: PNO141
R:Pankov, Y.A.; Karasev, V.S.
Biochimica 49, 1004-1018, 1994
A:Title: Luteinizing hormone of the sperm whale: amino acid sequence of reduced and car
A:Reference number: PNO141; MUID:84281133
A:Accession: PNO141
A:Molecule type: protein
A:Residues: 1-118 <PAN>

A>Note: article in Russian with English abstract
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone
F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.5%; Score 88; DB 2; Length 118;
Best Local Similarity 29.5%; Pred. No. 0.67;
Matches 36; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 80 CRELHYRFTVDGRCRSKAPVTELVCSGCG-CPARLLPNAIGRVKWRPNRNGDFRCIPDR 138
|||
DB 9 CRPINATLAQNZACPCVCIFFTTISICAGYCPSMVRVLPAL-----PPVPZPVCTYRQ 61
QY 139 YRAQRVOLL-CPGGAAPR-SRKVRVLVASCCKRLTFRHNSGLKDPETAPRQK-GRKP 195
|||
DB 62 LRFASIRLPGCCPPGVNPNVSPFVALSCHGCGPCRLSS-----SDGPGRAOPLACNRSF 114
QY 196 RP 197
||
DB 115 RP 116

RESULT 5

lutropin beta chain - minke whale

N:Alternate names: luteinizing hormone beta chain
C:Species: Balaenoptera acutirostrata (minke whale, lesser rorqual)
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 08-Dec-1995
C:Accession: PNO139
R:Karasev, V.S.; Pankov, Y.A.
Biochimica 50, 1972-1986, 1985
A:Title: Amino acid sequence of reduced and carboxymethylated alpha- and beta-subunit
A:Reference number: PNO138
A:Accession: PNO139
A:Molecule type: protein
A:Residues: 1-118 <KAR>
A>Note: article in Russian with English abstract
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone
F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.4%; Score 86; DB 2; Length 118;
Best Local Similarity 28.7%; Pred. No. 1;
Matches 35; Conservative 12; Mismatches 57; Indels 18; Gaps 6;

QY 80 CRELHYRFTVDGRCRSKAPVTELVCSGCG-CPARLLPNAIGRVKWRPNRNGDFRCIPDR 138
|||
DB 9 CRPINATLAQNZACPCVCIFFTTISICAGYCPSMVRVLPAL-----PPVPZPVCTYRQ 61
QY 139 YRAQRVOLL-CPGGAAPR-SRKVRVLVASCCKRLTFRHNSGLKDPETAPRQK-GRKP 195
|||
DB 62 LRFASIRLPGCCPPGVNPNVSPFVALSCHGCGPCRLSS-----SBCGPGRAZPLACBRSP 114
QY 196 RP 197
||
DB 115 RP 116

RESULT 6

luteinizing hormone beta-subunit - quail

C:Species: Coturnix coturnix (quail)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51242
R:Ando, H.; Ishii, S.
Gen. Comp. Endocrinol. 93, 357-366, 1994
A:Title: Molecular cloning of complementary deoxyribonucleic acids for the pituitary
urnix coturnix japonica).
A:Reference number: I51241; MUID:94252550

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:15 ; Search time 31.13 Seconds

(without alignments)
234.385 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 1167

Sequence: 1 MQLSLAPCLACLIVHAFAV.....KPRPRARAKANQAELENAY 213

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

rched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	95.5	8.2	159 1 LSHB_MELGA	P45646 meleagris g
2	88	7.5	118 1 LSHB_PHTCA	P23330 physeler ca
3	87.5	7.5	141 1 LSHB_TRIVU	O46482 trichosurus
4	86	7.4	118 1 LSHB_BALAC	P33088 balaenopter
5	83	7.1	474 1 KLF4_MOUSE	O60793 mus musculus
6	82.5	7.1	166 1 LSHB_COTJA	P45657 coturnix co
7	81.5	7.0	470 1 KLF4_HUMAN	O43474 homo sapien
8	81	6.9	170 1 PLGF_HUMAN	P49763 homo sapien
9	80	6.9	2142 1 BAT2_HUMAN	P48634 homo sapien
10	79.5	6.8	165 1 CSHB_HUMAN	P01233 homo sapien
11	77.5	6.6	138 1 LSHB_MACRU	O46483 macropus ru
12	77.5	6.6	188 1 VEGB_MOUSE	P49766 mus musculus
13	77.5	6.6	368 1 TGF4_MOUSE	O64280 mus musculus
14	77	6.6	1480 1 SLIT_DROME	P24014 drosophila
15	76.5	6.6	301 1 UL49_HSV11	P10233 herpes simp
16	76.5	6.6	361 1 LFTB_HUMAN	O75610 homo sapien
17	75.5	6.5	128 1 LSHB_STRCA	P80664 struthio ca
18	75.5	6.5	368 1 LFTB_MOUSE	P57785 mus musculus
19	75.5	6.5	1172 1 LMB3_HUMAN	O13751 homo sapien
20	75.5	6.5	2164 1 CCA4_MOUSE	P97445 mus musculus
21	75	6.4	141 1 LSHB_CERSI	O77835 ceratotheri
22	75	6.4	180 1 DAN_HUMAN	P41271 homo sapien
23	75	6.4	355 1 KLF2_HUMAN	O9Y5W3 homo sapien
24	75	6.4	379 1 CYR6_MOUSE	P18406 mus musculus
25	75	6.4	381 1 CYR6_HUMAN	O00622 homo sapien
26	75	6.4	628 1 MSLN_HUMAN	O13421 homo sapien
27	75	6.4	951 1 SPR8_HUMAN	O12872 homo sapien
28	74.5	6.4	322 1 YB1_MOUSE	P27817 mus musculus
29	74	6.3	174 1 BAR1_CHITE	P02849 chironomus
30	74	6.3	366 1 TGF4_HUMAN	O00292 homo sapien
31	74	6.3	563 1 MUC5_BOVIN	P98091 bos taurus
32	73.5	6.3	324 1 YB1_HUMAN	P16991 homo sapien
33	73.5	6.3	1436 1 WC11_BOVIN	P30205 bos taurus

34	73.5	6.3	3133 1 HMCT_BOMO	P98092 bombyx mori
35	73	6.3	305 1 YB3_XENLA	O00436 xenopus lae
36	73	6.3	511 1 NEX3_MOUSE	O9-025 mus musculus
37	73	6.3	581 1 PRKR_BOVIN	O28172 bos taurus
38	73	6.3	644 1 VP4_BRY13	P33429 bluetongue
39	73	6.3	1955 1 AGRI_CHICK	P31696 gallus galli
40	72.5	6.2	551 1 YOLL_MYCTU	P71732 mycobacteri
41	72.5	6.2	1056 1 MUC5_HUMAN	P98088 homo sapien
42	72	6.2	644 1 VP4_BRY11	P33428 bluetongue
43	72	6.2	888 1 KLRK_MOUSE	P08823 mus musculus
44	72	6.2	975 1 CDP_CANFA	P39881 canis fam11
45	72	6.2	1268 1 PCGN_MOUSE	P55066 mus musculus

ALIGNMENTS

RESULT ID	LSHB_MELGA	STANDARD:	PRT:	159 AA.
AC	P45646:			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).			
DE	LHB.			
GN	Meleagris gallipavo (Common turkey).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.			
OX	NCBI_TaxID=9103;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pituitary;			
RX	MEDLINE=95290073; PubMed=7772235;			
RA	You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;			
RT	"Sequence analysis of the turkey LH beta subunit and its regulation by gonadotrophin-releasing hormone and prolactin in cultured pituitary cells.";			
RT	J. Mol. Endocrinol. 14:117-129(1995).			
RL	-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.			
CC	-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN.			
CC	LUTROPIN, FOLLITROPIN AND GONADOTROPIN.			
CC	-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sid.ch).			
CC	-----			
DR	EMBL: L35519; AAA74125.1; ALT_INIT.			
DR	HSSP: P01233; IHRP.			
DR	InterPro: IPR000359; -.			
DR	InterPro: IPR001545; -.			
DR	InterPro: IPR002400; -.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	PRINTS: PR00438; GFCYSKNOT.			
DR	PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.			
DR	PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.			
KW	Hormone; Signal; Glycoprotein.			
FT	SIGNAL 1			POTENTIAL.
FT	CHAIN 40			LUTROPIN BETA CHAIN.
FT	DISULFID 48			BY SIMILARITY.
FT	DISULFID 62			BY SIMILARITY.
FT	DISULFID 65			BY SIMILARITY.
FT	DISULFID 73			BY SIMILARITY.
FT	DISULFID 77			BY SIMILARITY.

FT DISULFID 132 139 BY SIMILARITY.
 FT CARBOHYD 52 52 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 159 AA: 16265 MM: 52850C8C879653C6 CRC64;

Query Match 8.2%; Score 95.5; DB 1; Length 159;
 Best Local Similarity 27.4%; Pred. No. 0.047;
 Matches 34; Conservative 11; Mismatches 58; Indels 21; Gaps 5;

QY 62 GGRPHHPDTRKDYSEYSCRELHYTFYVDGCRSAKAPVTELYCSGGCPARLLPNAIGR 121
 DB 43 GGRPP-----CRPIWTVAVERDECPQCAVTTTACGXCRT-----EPVYR 85
 QY 122 VKMMRPNGPDFRCIPDRYARQVOLL-CPGGAAPRSRKVRLVASCCKRRLTRFHNSLEK 180
 DB 86 SPAGRP--PQSSCTGALATERNALMGCPITGSDPRV-LLPVALSCRCARCPITATSDCTVQ 142
 QY 181 DEGP 184
 DB 143 GLGP 146

RESULT 2
 LSHB_PHYCA STANDARD; PRT: 118 AA.

AC P25330;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LUTROPIN BETA CHAIN (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA)
 GN LHB.
 OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
 OC Physetereidae; Physeter.
 OX NCBI_TaxID=9755;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=87032654; PubMed=3771098;
 RA Pankov Y.A., Karasev V.S.;
 RT "Primary structure of sperm whale luteinizing hormone.";
 RL Int. J. Pept. Protein Res. 28:124-129(1986).
 RN (2)
 RP SEQUENCE.
 RX MEDLINE=84281133; PubMed=6466737;
 RA Pankov Y.A., Karasev V.S.;
 RT "Luteinizing hormone of the sperm whale. Amino acid sequences of
 reduced and carboxymethylated beta-subunits.";
 RL Biochimia 49:1004-1018(1984).

-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 FAMILY.

CC PIR: PNCI41; PNCI41.
 DR HSSP: P01233; IHRP.
 DR InterPro: IPR000359; -;
 DR InterPro: IPR001545; -;
 DR InterPro: IPR002400; -;
 DR Pfam: PF00007; Cys_knot.1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 9 57 BY SIMILARITY.
 FT DISULFID 23 72 BY SIMILARITY.
 FT DISULFID 26 110 BY SIMILARITY.
 FT DISULFID 34 88 BY SIMILARITY.
 FT DISULFID 38 90 BY SIMILARITY.
 FT DISULFID 93 100 BY SIMILARITY.

FT CARBOHYD 13 13 N-LINKED (GLCNAC...)
 SQ SEQUENCE 118 AA: 12412 MM: 81177A56382F15E7 CRC64;

Query Match 7.5%; Score 88; DB 1; Length 118;
 Best Local Similarity 29.5%; Pred. No. 0.17;
 Matches 36; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 80 CRELHYTFEYVDGCRSAKAPVTELYCSGQC-GPARLLPNAIGRKVMRPNPGRFCIPDR 138
 DB 9 CRPINATLAQNZACPVCITFTTISICAGYCPSMWRVLPAL-----PPVZPVCTYRQ 61
 QY 139 YRAQVOLL-CPGGAAPR-SRKVRLVASCCKRRLTRFHNSLEKDFGPETAPRQK-GRKP 195
 DB 62 LRFASTRPGCPPEVNPVSPFVALSCHCGCRLLS-----SDCGRGAQPLACNRSP 114
 QY 196 RP 197
 DB 115 RP 116

RESULT 3
 LSHB_TRIUV STANDARD; PRT: 141 AA.

AC O46482;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA)
 GN LHB.
 OS Trichosurus vulpecula (Brush-tailed possum).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
 OX NCBI_TaxID=9337;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=98345424; PubMed=9680384;
 RA Harrison G.A., Deane E.M., Cooper D.W.;
 RT "cDNA cloning of luteinizing hormone subunits from brushtail possum
 and red kangaroo.";
 RL Mamm. Genome 9:638-642(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Lawrence S.B., McNatty K.P., Fidler A.E.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed, usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF017448; AAC96019.1; -;
 DR EMBL: AF090388; AAC63526.1; -;
 DR HSSP: P01233; IHRP.
 DR InterPro: IPR000359; -;
 DR InterPro: IPR001545; -;
 DR Pfam: PF00007; Cys_knot.1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 KW Hormone; Signal; Glycoprotein.
 FT SIGNAL 1 22 BY SIMILARITY.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:00:36 : Search time 89.34 Seconds

(without alignments)
315.435 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 1167
Sequence: 1 MGLSLAPCLACLVHAFVA.....KPRPRARGAKANOAELENAY 213

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

rchd: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	256.5	22.0	134	Q9Y3U3	Q9Y3U3 homo sapien
2	96	8.2	272	Q9PWB0	Q9PWB0 gallus gall
3	94	8.1	272	Q9PWB2	Q9PWB2 gallus gall
4	93.5	8.0	1114	Q9JRW7	Q9JRW7 mus musculu
5	92	7.9	712	Q9Y557	Q9Y557 homo sapien
6	92	7.9	720	Q9NY23	Q9NY23 homo sapien
7	90.5	7.8	493	Q9UEW3	Q9UEW3 homo sapien
8	88	7.5	168	Q9H772	Q9H772 homo sapien
9	87	7.5	184	Q73755	Q73755 gallus gall
10	87	7.5	381	Q88995	Q88995 mus musculu
11	87	7.5	988	Q97867	Q97867 sus scrofa
12	87	7.5	1537	Q9WTW6	Q9WTW6 mus musculu
13	87	7.5	1581	Q9R0B6	Q9R0B6 mus musculu
14	86.5	7.4	493	Q9S376	Q9S376 homo sapien
15	86	7.4	925	Q9F470	Q9F470 porphyromon
16	85.5	7.3	481	Q9HBZ6	Q9HBZ6 homo sapien
17	85.5	7.3	791	Q70368	Q70368 rattus norv
18	83	7.1	168	Q88273	Q88273 mus musculu
19	83	7.1	483	Q9R255	Q9R255 mus musculu

20	82.5	7.1	184	4	Q60565	Q60565 homo sapien
21	82	7.0	1035	5	Q9NEG3	Q9NEG3 drosophila
22	81.5	7.0	470	4	Q9UNP3	Q9UNP3 homo sapien
23	81	6.9	2157	4	Q95875	Q95875 homo sapien
24	81	6.9	3163	2	Q92373	Q92373 yersinia pe
25	80.5	6.9	215	8	Q9WFB6	Q9WFB6 beta vulgar
26	80.5	6.9	1081	4	Q76065	Q76065 homo sapien
27	80.5	6.9	1704	5	Q94446	Q94446 chironomus
28	80	6.9	379	11	Q9ES72	Q9ES72 rattus norv
29	80	6.9	1096	10	Q9XE24	Q9XE24 oryza sativ
30	80	6.9	3161	2	Q54511	Q54511 yersinia en
31	78.5	6.7	492	11	Q9Z1K6	Q9Z1K6 mus musculu
32	78.5	6.7	761	3	Q9HEA3	Q9HEA3 neurospora
33	78.5	6.7	2098	5	Q25757	Q25757 plasmodium
34	78.5	6.7	2114	5	Q97267	Q97267 plasmodium
35	78	6.7	343	11	Q61406	Q61406 mus musculu
36	77.5	6.6	184	11	Q35793	Q35793 rattus norv
37	77.5	6.6	184	11	Q70326	Q70326 mus musculu
38	77.5	6.6	270	13	P70041	P70041 xenopus lae
39	77.5	6.6	482	14	Q71654	Q71654 human immun
40	77.5	6.6	1637	6	Q9XSV8	Q9XSV8 bos taurus
41	77	6.6	267	4	Q95813	Q95813 homo sapien
42	77	6.6	355	4	Q9UKR6	Q9UKR6 homo sapien
43	77	6.6	384	10	Q9FSQ3	Q9FSQ3 oryza sativ
44	77	6.6	385	5	Q27008	Q27008 toxoplasma
45	77	6.6	826	10	Q9FU12	Q9FU12 oryza sativ

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	134 AA.
Q9Y3U3	Q9Y3U3			
AC	Q9Y3U3:			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	HYPOHETICAL 15.3 KDA PROTEIN (FRAGMENT).			
GN	DKEZP564D206.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RA	Wambolt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL050024; CAB43243.1; -			
DR	InterPro: IPR000359; -			
DR	PROSITE: PS01225; CTCK_2; 1.			
KW	Hypothetical protein.			
FT	NON_TER			
SQ	SEQUENCE 134 AA; 15324 MW; A0B7A8650D2E6F1 CRC64;			

Query Match 22.0%; Score 256.5; DB 4; Length 134;
Best Local Similarity 41.9%; Pred. No. 1.8e-18;
Matches 52; Conservative 24; Mismatches 43; Indels 5; Gaps 3;

QY	80	CRELHYTRFYVDGFCRSAPVTELVCSGCGPARLLPNAIG--RVKMW-RPNGPDFRCI	135
DB	3	CRELRSKRYISDGGCTSIPLKELVCAEGEDLLPVLPNNWIGGVTGRYMSRRSSQEMRCV	62
QY	136	PDRTRARVOLLICGGAAPRSKRYRLVASCKRKLRFPHNQSLEKDPGPTARPGGRGP	195
DB	63	NDKTRTRQRIQLQCODGST-RYKITYVTACKCKRYTRQHNSSHNFSMSAPKPVQHRE	121
QY	196	RPR 199	
DB	122	RKRA 125	

```
RESULT 2
O9PMB0 PRELIMINARY; PRT; 272 AA.
AC O9PMB0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
GN CER.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart."
Query Match 8.2%; Score 96; DB 13; Length 272;
Best Local Similarity 22.2%; Pred. No. 0.065;
Matches 30; Conservative 22; Mismatches 57; Indels 26; Gaps 5;
OY 45 EPPQELN-----NOTMNAENGRRPHHPYDTKRYSEYSCRELHYTFYDGCPSAKPV 100
DB 129 EPPYRKDAKKFWHDHMLKRNASSEVVLPIKTNEHOCRTLPFSQVAHSCSEKV-IV 187
OY 101 TELVSGGCGPARLLPNAIGRVKWRPNPGPDR-----CIPDRYRAORVOLLCPGGA 153
DB 188 ONNLGCGKCSS-----FHVGPDDRLLYTFCSKCLPTKFSMKHDLNCT-SSV 233
OY 154 PRSRKRVLVASCCK 168
DB 234 PVVKKVMIVEECNCE 248
RESULT 3
O9PUK2 PRELIMINARY; PRT; 272 AA.
AC O9PUK2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CARONTE.
GN CAR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez Esteban C., Capdevila J., Economides A.N., Pascual J.,
RA Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-like protein, mediates the establishment of
RT embryonic left-right asymmetry."
RT Nature 0:0-0(1999).
RL EMBL; AF179484; AAD55581.1; -.
DR InterPro; IPR000359; -.
DR InterPro; IPR001839; -.
DR ProDom; PD000357; -. 1.
DR PROSITE; PS01225; CTCK_2; 1.
```

```
DR SMART; SM00041; CT; 1.
SQ SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;
Query Match 8.1%; Score 94; DB 13; Length 272;
Best Local Similarity 22.2%; Pred. No. 0.1;
Matches 30; Conservative 21; Mismatches 58; Indels 26; Gaps 5;
OY 45 EPPQELN-----NOTMNAENGRRPHHPYDTKRYSEYSCRELHYTFYDGCPSAKPV 100
DB 129 EPPYRKDAKKFWHDHMLKRNASSEVVLPIKTNEHOCRTLPFSQVAHSCSEKV-IV 187
OY 101 TELVSGGCGPARLLPNAIGRVKWRPNPGPDR-----CIPDRYRAORVOLLCPGGA 153
DB 188 ONNLGCGKCSS-----FHVGPDDRLLYTFCSKCLPTKFSMKHFDLNC-SSV 233
OY 154 PRSRKRVLVASCCK 168
DB 234 PVVKKVMIVEECNCE 248
RESULT 4
O9UKW7 PRELIMINARY; PRT; 1114 AA.
AC O9UKW7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CYSTEINE-RICH PROTEIN NFX-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ariotta P., Jenkins N.A., Jay G., Ono S.J.;
RT "Isolation of a Full-Length Murine NFX-1 cDNA, Its Chromosomal
RT Location and Developmental Expression."
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223576; AAF34700.1; -.
DR InterPro; IPR000967; -.
DR InterPro; IPR001374; -.
DR InterPro; IPR001841; -.
DR InterPro; IPR001965; -.
DR Pfam; PF01422; zf-NF-X1; 8.
DR Pfam; PF01424; R3H; 1.
SQ SEQUENCE 1114 AA; 123877 MW; 857C90260D28EAB CRC64;
Query Match 8.0%; Score 93.5; DB 11; Length 1114;
Best Local Similarity 23.6%; Pred. No. 0.5;
Matches 48; Conservative 26; Mismatches 90; Indels 39; Gaps 9;
OY 21 VESGQWQAFKND-----ATEIIRGLREYEPPEQLENNQTMNRAENGRR-PHHPIYDTKDV 75
DB 98 LONOPWQAKARNEKHONNRKKAQGLSEOTSSTSSLE---SVARSSESGNPREHSPSE-- 152
OY 76 SEYSCRELHYTFYDGCPSAKPVTELVSGGCGPARLLPNAIGRVK--WRRPNPGPDR 133
DB 153 -----KEVVIADPRGAKPKKAAQLIYNGRG--PKAKGRARSMGMRMSKSE 198
OY 134 C-IPDRYRAORVOLLCPGGAAPRSRKRVLVASCCKRLTPRHNSGLKDFGPETARPQK 192
DB 199 DEIDPPMWFPTLTQIASCRKP-----VDDPCVCRRNEORRYRQKRPPLGSGRARPRG 252
OY 193 RKP-----RRRARGAKANQAEI 209
DB 253 RNPKQESQRHINAGPKTNMSPI 275
RESULT 5
O9Y557 PRELIMINARY; PRT; 712 AA.
ID O9Y557
```


XX Claim 6; Page 125-126; 162pp; English.
PS
CC This shows a rat transforming growth factor-beta (TGF-beta) binding
CC protein designated rBEER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patient's own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.
XX
SQ Sequence 213 AA;
Query Match 100.0%; Score 1167; DB 21; Length 213;
Best Local Similarity 100.0%; Pred. No. 6,7e-102;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MQLSLAPCLACLLVHAFAVAVESGQMAFKNDATETIIPGLREYEPPEOLENNQTNRAE 60
Db 1 mqlslapclacilivhaafavessqgwfkdntellipglreyeppeqlennqtnrae 60
QY 61 NGRPRPHRYDTRDSEYSCRELHATRTFTDGPGRSAKPVTELVCSGCGPARLPNAIG 120
Db 61 ngrprphrydtkdvseyrcrelhytrftvdgpcrsakpvtelvcsgcgparlpnaig 120
QY 121 RYKWMRPNGPDRCPDRYRAQROVOLLCPGGAAPRSKRVLVASCKKRLTRFHNSSELK 180
Db 121 rykwmrpnpgdrctipdryraqrvqlcp9gaaprskrvllvasckkrltrfhngselk 180
QY 181 DFGPETARPOKGRKPRPARAGAKANOAELENAY 213
Db 181 dfgpetarpqkgrkprparagakanqaelenay 213
RESULT 2
AA96432 standard; Protein; 211 AA.
XX
AC AAY96432;
XX
DT 12-SEP-2000 (first entry)
XX
XX Murine TGF-beta binding protein (BEER).
XX
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX
OS Mus musculus.
XX
PN MO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
PA (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
PI Van Ness J, Winkler DG;
XX
XX WPI; 2000-412321/35.
XX
DR N-PSDB; AAA29058.

XX Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
PS
XX Claim 5; Page 124; 162pp; English.
XX
CC This shows a murine transforming growth factor-beta (TGF-beta)
CC binding protein designated mBEER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patient's own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.
XX
SQ Sequence 211 AA;
Query Match 93.7%; Score 1094; DB 21; Length 211;
Best Local Similarity 93.9%; Pred. No. 4.7e-95;
Matches 200; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
1 MQLSLAPCLACLLVHAFAVAVESGQMAFKNDATETIIPGLREYEPPEOLENNQTNRAE 60
Db 1 mqlslapclacilivhaafavessqgwfkdntellipglreyeppeolennqtnrae 58
QY 61 NGRPRPHRYDTRDSEYSCRELHATRTFTDGPGRSAKPVTELVCSGCGPARLPNAIG 120
Db 59 ngrprphrydtkdvseyrcrelhytrftvdgpcrsakpvtelvcsgcgparlpnaig 118
QY 121 RYKWMRPNGPDRCPDRYRAQROVOLLCPGGAAPRSKRVLVASCKKRLTRFHNSSELK 180
Db 119 rykwmrpnpgdrctipdryraqrvqlcp9gaaprskrvllvasckkrltrfhngselk 178
QY 181 DFGPETARPOKGRKPRPARAGAKANOAELENAY 213
Db 179 dfgpetarpqkgrkprparagakanqaelenay 211
RESULT 3
AA96431 standard; Protein; 213 AA.
XX
AC AAY96431;
XX
DT 12-SEP-2000 (first entry)
XX
XX Vervet TGF-beta binding protein (BEER).
XX
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX
OS Cercopithecus pygerythrus.
XX
PN MO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
PA (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:15 ; Search time 31.13 Seconds
(without alignments)
193.671 Million cell updates/sec

Title: US-09-668-021-16

Perfect score: 967
Sequence: 1 NDATETIPELGEYEPLEL.....RPGTKRLRPANGTKASRA 176

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

riched: 93435 segs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	99.5	10.3	159	1	LSHB_MELGA
2	83.5	8.6	166	1	LSHB_COTJA
3	82.5	8.5	141	1	LSHB_TRIVU
4	81	8.4	455	1	RADA_TREPA
5	80	8.3	118	1	LSHB_PHYCA
6	79.5	8.2	1480	1	SLIT_DROME
7	79	8.2	470	1	KLFA_HUMAN
8	79	8.2	682	1	SNK_MOUSE
9	79	8.2	682	1	SNK_RAT
10	79	8.2	685	1	SNK_HUMAN
11	78	8.1	118	1	LSHB_BAIAC
12	77.5	8.0	123	1	LSHB_STRCA
13	77	8.0	243	1	TOMB_ENTAE
14	76.5	7.9	628	1	MSLN_HUMAN
15	75.5	7.8	141	1	LSLN_CERSI
16	75.5	7.8	474	1	KLFA_MOUSE
17	75	7.8	165	1	CGHB_PARAN
18	74	7.7	398	1	MOBI_XENLA
19	74	7.7	544	1	H15_DROME
20	73.5	7.6	870	1	BCAL_HUMAN
21	73	7.5	138	1	LSHB_MACRU
22	73	7.5	256	1	OBFL_HUMAN
23	72	7.4	143	1	LSHB_FELCA
24	72	7.4	165	1	CGHB_HUMAN
25	71.5	7.4	450	1	A2AA_MOUSE
26	70.5	7.3	1134	1	TRE1_MOUSE
27	70.5	7.3	1496	1	CA25_HUMAN
28	70	7.2	178	1	DAN_RAT
29	70	7.2	891	1	MAY1_SCHCO
30	70	7.2	972	1	SVL_MYCLE
31	70	7.2	2182	1	CABI_RAT
32	70	7.2	2261	1	ABCI_HUMAN
33	69.5	7.2	200	1	RR4_GUTHH

34	69.5	7.2	356	1	ENV_FRSPB
35	69.5	7.2	998	1	EPB3_HUMAN
36	69.5	7.2	1436	1	WC11_BOVIN
37	69	7.1	180	1	DAN_HUMAN
38	69	7.1	248	1	HUTC_PSEPU
39	69	7.1	355	1	KLFE2_HUMAN
40	69	7.1	543	1	PKAL_STRCO
41	69	7.1	837	1	MUCL_RAT
42	69	7.1	1093	1	SM5B_MOUSE
43	69	7.1	1746	1	TENA_PIG
44	68.5	7.1	1056	1	MUC5_HUMAN
45	68	7.0	138	1	MUC5_CANFA

ALIGNMENTS

```

RESULT 1
LSHB_MELGA STANDARD, PRT, 159 AA.
ID LSHB_MELGA
AC P45646;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-
DE BETA) (LSH-B) (LH-B).
GN LHB.
OS Melagris gallinavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_Taxid=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutary;
RX MEDLINE=95290073; PubMed=7772235;
RA You S., Foster L.K., Sliby J.L., el Halawani M.E., Foster D.N.;
RT "Sequence analysis of the turkey LH beta subunit and its regulation
RT by gonadotropin-releasing hormone and prolactin in cultured
RT pituitary cells.";
RT J. Mol. Endocrinol. 14:117-129(1995).
RL -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
CC -!- THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERES BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: L35519; AAA74125.1; AL1_INT1.
CC HSSP: P01233; IHRP.
CC InterPro: IPR000359; -.
CC InterPro: IPR001545; -.
CC Pfam: PF00007; Cys_knot. 1.
CC PRINTS: PR00438; GFCYSKNOT.
CC PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
CC Hormone; Signal; glycoprotein.
CC KW SIGNAL; 1
CC CHAIN 40
CC FT DISULFID 48 159
CC FT DISULFID 62 111
CC FT DISULFID 65 149
CC FT DISULFID 73 127
CC FT DISULFID 77 129
CC FT BY SIMILARITY.
CC FT BY SIMILARITY.
CC FT BY SIMILARITY.
CC FT BY SIMILARITY.

```

FT DISULFID 132 139 BY SIMILARITY
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 159 AA; 16285 MW; 52850C8C879653C6 CRC64;

Query Match 10.3%; Score 99.5; DB 1; Length 159;
 Best Local Similarity 27.4%; Pred. No. 0.008;
 Matches 34; Conservative 11; Mismatches 58; Indels 21; Gaps 5;

QY 31 GGRPHHPEFTKDASEYSCRELHFTTRYVDGPCRSKAPYTELVCSGCGCPARLLPNAIGR 90
 DB 43 GGRPP-----CRPINTVAVEKEDECPQCMAMVTTTACGGCYCTREPV-----EPVYR 85
 QY 91 GKWRPSGDPFCIPDRYRAQRVOLL-CPGGAAPRAKRVLVASCKCKRLTRFHNSLKL 149
 DB 86 SPLGRP--PQSSCTYGALRYERMALMGCPIGSDPRV-LLPVALSCRCARCPATSDCTVQ 142
 QY 150 DFGP 153
 DB 143 GLGP 146
 ID LSHB_CONJUA STANDARD; PRT; 166 AA.
 AC P45657;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-
 DE BETA) (LSH-B) (LH-B).
 GN LHB.
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix;
 OC NCBI_TaxID=93934;
 RN [1]
 RX MEDLINE=94252550; PubMed=7515015;
 RA Ando H., Ishii S.;
 RT "Molecular cloning of complementary deoxyribonucleic acids for the
 RT pituitary glycoprotein hormone alpha-subunit and luteinizing hormone
 RT beta-subunit precursor molecules of Japanese quail (Coturnix coturnix
 RT japonica)."
 RL Gen. Comp. Endocrinol. 93:357-368(1994).
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 CC EMBL: S70834; AAB30867.1; -;
 DR HSSP: P01233; IHRP.
 DR InterPro: IPR000359; -;
 DR InterPro: IPR001545; -;
 DR InterPro: IPR002400; -;
 DR Pfam: PF00007; Cys_knot; 1.
 DR PRINTS: PR00438; GEGYKNOF.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 DR Hormone; Signal; 1 21 POTENTIAL.
 FT SIGNAL 1 21

FT CHAIN 22 166 LUTROPIN BETA CHAIN.
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 56 104 BY SIMILARITY.
 FT DISULFID 70 119 BY SIMILARITY.
 FT DISULFID 73 157 BY SIMILARITY.
 FT DISULFID 81 135 BY SIMILARITY.
 FT DISULFID 85 137 BY SIMILARITY.
 FT DISULFID 140 147 BY SIMILARITY.
 SO SEQUENCE 166 AA; 17030 MW; 6BF293BEC3C5FAC7 CRC64;

Query Match 8.6%; Score 83.5; DB 1; Length 166;
 Best Local Similarity 24.8%; Pred. No. 0.29;
 Matches 31; Conservative 13; Mismatches 56; Indels 25; Gaps 5;

QY 32 GRRPHHPEFTKDASEYSCRELHFTTRYVDGPCRSKAPYTELVCSGCGCPARLLPNAIGR 91
 DB 52 GRPP-----CRPINTVAVEKEDECPQCMAMVTTTACGGCYCTREPV-----EPVYR 91
 QY 92 KWRPSG--PDFCIPDRYRAQRVOLL-CPGGAAPRAKRVLVASCKCKRLTRFHNSLKL 148
 DB 92 YRSLGPPPPQSSCTYGALRYERMALMGCPIGSDPRV-LLPVALSCRCARCPATSDCTVQ 149
 QY 149 KDFGP 153
 DB 150 QGLGP 154
 ID LSHB_TRIYU STANDARD; PRT; 141 AA.
 AC O46482;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-
 DE BETA) (LSH-B) (LH-B).
 GN LHB.
 OS Trichosurus vulpecula (Brush-tailed possum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
 OC NCBI_TaxID=9337;
 RN [1]
 RX MEDLINE=9680384; PubMed=9680384;
 RA Harrison G.A., Deane E.M., Cooper D.W.;
 RT "CDNA cloning of luteinizing hormone subunits from brushtail possum
 RT and red kangaroo."
 RL Mamm. Genome 9:638-642(1998).
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 CC EMBL: AF017448; AAC96019.1; -;
 DR HSSP: AF090388; AAC63526.1; -;
 DR HSSP: P01233; IHRP.


```

RESULT 2
ID Q9PWB0 PRELIMINARY: PRT: 272 AA.
AC Q9PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
GN CER.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gaidiner A., Lassar A.B., Mercola M., Stern C.D.,
Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart.";
CURT. Biol. 9:931-938(1999).
EMBL: AF139721; AAD51610.1; -.
InterPro: IPR000359; -.
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT; 1.
SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match
Best Local Similarity 10.5%; Score 102; DB 13; Length 272;
Matches 26; Conservative 18; Mismatches 41; Indels 22; Gaps 4;

Y 38 PFETKDASEYSCRELHFTRYVTDGPGRSAPVTELVCSGCGPARLLPNAIGRGKMMRPS 97
D 157 PIKTNEMHOETCRTPPESSVAHSECEKV-IVONNICFGKCSS-----FHVP 202
Y 98 GPDR-----CIPDRYRAORVQLCPGGAAPRAKRYLVASCCK 137
D 203 GPDDRLLTFGSKCLPTKFSMKHDLNCT-SSVPVKKYIVAEBCNCE 248

RESULT 3
ID Q9PUBK2 PRELIMINARY: PRT: 272 AA.
AC Q9PUBK2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CAROMTE.
GN CAR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX Rodriguez Esteban C., Capdevila J., Economides A.N., Pascual J.,
Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-like protein, mediates the establishment of
RT embryonic left-right asymmetry.";
NL Nature 0:0-0(1999).
DR EMBL: AF179484; AAD55581.1; -.
DR InterPro: IPR000359; -.
DR InterPro: IPR001839; -.
DR PRODOM: PD000357; -.
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT; 1.
SQ SEQUENCE 272 AA; 31201 MW; 00B9729715771BC2 CRC64;

```

```

Query Match
Best Local Similarity 10.3%; Score 100; DB 13; Length 272;
Matches 26; Conservative 17; Mismatches 42; Indels 22; Gaps 4;

Y 38 PFETKDASEYSCRELHFTRYVTDGPGRSAPVTELVCSGCGPARLLPNAIGRGKMMRPS 97
D 157 PIKTNEMHOETCRTPPESSVAHSECEKV-IVONNICFGKCSS-----FHVP 202
Y 98 GPDR-----CIPDRYRAORVQLCPGGAAPRAKRYLVASCCK 137
D 203 GPDDRLLTFGSKCLPTKFSMKHDLNCT-SSVPVKKYIVAEBCNCE 248

RESULT 4
ID Q9Y557 PRELIMINARY: PRT: 712 AA.
AC Q9Y557;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE DJ116311.3 (NOVEL PROTEIN SIMILAR TO MOUSE B99) (FRAGMENT).
GN DJ116311.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RT Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL031588; CAB38415.1; -.
FT NON-TER
SQ SEQUENCE 712 AA; 75685 MW; ECBD2133EF72F32 CRC64;

Query Match
Best Local Similarity 9.6%; Score 93; DB 4; Length 712;
Matches 34; Conservative 21; Mismatches 37; Indels 60; Gaps 6;

Y 62 PCRSAPVTELVCSGCGPARILP-----NAIGRGKMMRSGDFPCIPDRYRAORV----- 113
D 333 PANSRPLSNISKSGRGAPMLRPALPACPGVASSM-----QAKRVVSE 377
Y 114 ---QLCPGGAP-----RAKRYLVASC---KCKRLT 140
D 378 LAEOLTRAPPSASPTOPQTPGCGGOWLNSSCAWESSQLNKRIRRDSCINXKWKVP 437
Y 141 RPHNSELXDF---GPEARAPOTGRKLRPRA 168
D 438 TPTNOFKIPKFSIDSPDSSITPKLSRAORFOS 469

RESULT 5
ID Q9NY23 PRELIMINARY: PRT: 720 AA.
AC Q9NY23;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE B99.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Monte M., Collavin L., Lazarevic D., Utrera R., Schneider C.;
RT "Cloning, chromosome mapping and functional characterization of a
RT human homolog gene of murine B99.";
NL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF223408; AAF31459.1; -.
SQ SEQUENCE 720 AA; 76628 MW; D4E852CE4BE85BBF CRC64;

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:43 ; Search time 80.28 Seconds

(without alignments)
132.908 Million cell updates/sec

Title: us-09-668-021-16

Perfect score: 967
Sequence: 1 MDATETPELGEFPEPEPEL.....RPTGKRLPRARGTAKSA 176

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

riched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_0601.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	967	100.0	176	21	AAV96434
2	920.5	95.2	213	21	AAV96431
3	901.5	93.2	213	21	AAV96429
4	901.5	93.2	213	22	AAV96430
5	892.5	92.3	213	21	AAV96436
6	892.5	92.3	213	21	AAV96433
7	861.5	89.1	213	21	AAV96432
8	857.5	83.6	367	21	AAV96432
9	808.5	34.7	206	21	AAV96432
10	808.5	34.7	206	21	AAV96432
11	336				AAV75981

12	336	34.7	206	21	AAV76031	Murine skin cell p
13	336	34.7	206	22	AAV55920	Skin cell protein,
14	336	34.7	206	22	AAV55970	skin cell protein,
15	333	34.4	206	18	AAV09408	Human small CCN-11
16	333	34.4	206	19	AAV58704	Human small CCN-11
17	333	34.4	206	20	AAV95711	Homo sapiens fetal
18	333	34.4	206	21	AAV10233	Human adult retina
19	108.5	11.2	116	18	AAW27654	Secreted protein A
20	108.5	11.2	116	18	AAW44090	Human secreted pro
21	103.5	10.7	102	20	AAV03225	Amino acid sequenc
22	103.5	10.7	102	21	AAV10277	Human fetal kidney
23	96	9.9	50	20	AAV12009	Human 5' EST seque
24	93	9.6	712	21	AAV43147	Human ORF ORF2911
25	87	9.0	168	21	AAV84014	Amino acid sequenc
26	86	8.9	168	21	AAV51132	Murine cerebral ne
27	85.5	8.8	320	20	AAV04748	Mycobacterium spec
28	85	8.8	184	19	AAW29727	DAN and b57 protei
29	83.5	8.6	1428	21	AAV97033	Caspase 8-interact
30	82.5	8.5	145	12	AAV15113	hCG/ELH chimera, E
31	80.5	8.3	212	20	AAV43278	Human CG beta subu
32	79.5	8.2	1480	13	AAV25079	Drosophila SLIT pr
33	79	8.2	212	20	AAV43282	hCG/hTSH beta subu
34	79	8.2	246	21	AAV78333	Herpes simplex vir
35	79	8.2	470	20	AAV24314	Human RPLF protein
36	79	8.2	470	21	AAV13773	Human RPLF protein
37	79	8.2	470	22	AAV48018	Human RPLF protein
38	79	8.2	479	20	AAV24315	Human RPLF protein
39	79	8.2	479	21	AAV13774	Human RPLF protein
40	79	8.2	685	20	AAV00915	Disease associated
41	79	8.2	685	20	AAV88432	Human CG beta subu
42	78.5	8.1	204	20	AAV43275	hCG/ELH chimera, E
43	77.5	8.0	145	12	AAV15109	Chorionic gonadotr
44	77.5	8.0	145	18	AAV27680	Human DM protein
45	77.5	8.0	184	20	AAV42172	

ALIGNMENTS

RESULT 1	
ID	AAV96434 standard: Protein: 176 AA.
XX	AAV96434:
AC	12-SEP-2000 (first entry)
XX	
DT	
XX	
DE	Bovine TGF-beta binding protein (BEER).
XX	
XX	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW	BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX	
OS	Bos taurus.
XX	
PN	WO200032773-A1.
XX	
PD	08-JUN-2000.
XX	
PF	24-NOV-1999; 99WO-US27990.
XX	
PR	27-NOV-1998; 98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
XX	
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
PI	Van Ness J, Winkler DG;
XX	
DR	WPI, 2000-412321/35.
XX	
DR	N-PDOB; AAA29060.
XX	
PT	Nucleic acids (I) encoding a transforming growth factor beta binding
PT	protein, useful for identifying agents for treating osteopenia,
PT	osteoporosis and fractures

XX Claim 7: Page 127: 162pp: English.

Query Match	100.0%	Score 967;	DB 21;	Length 176;
Best Local Similarity	100.0%	Pred. No. 1.5e-87;		
Matches 176; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Oy	1	NDANEI1PELGEYEPYEP1ELNNKTKMNAENGRPHNFPEFKDASXSCHELIHFTTRVTD	60
Db	1	ndate1p1elgeyp1p1elnnk tcmnaengrphnp fekdasey scret1hfttrytd	60
Oy	61	GPCRSARKVTE1YCSGGCGPARLLPNAIGRCMKWRPSCGPEDRCIPDXYRAORVOLLCPG	120
Db	61	gpcrsarkvte1ycsqgqparllpna1grgkwwrpsgpdtrc1pdxryaqrvll1oppg	120
Oy	121	AAPRARKRVLYASCKRRLTFHMQSELKDPGPPAARPOGRKIRPRARAGTKASRA	176
Db	121	aaprarkrvlyasckrcltrfhnqselkdpgpaarpqgrk1rpraragtkasra	176

XX Claim 4; Page 122-123; 162pp; English.

XX

XX This shows a vervet transforming growth factor-beta (TGF-beta)

CC binding protein designated vBER. The cDNA and protein may be used for

CC prevention, treatment and diagnosis of diseases associated with

CC inappropriate BEER expression. For example, they may be used to treat

CC disorders associated with decreased TGF-beta BP expression. The cDNA or

CC vectors may be administered to treat diseases by rectifying mutations or

CC deletions in a patient's genome that affect the activity of BEER by

CC expressing inactive proteins or to supplement the patients own production

CC of BEER polypeptides. The nucleic acids may be used for recombinant

CC production of BEER, gene therapy, antisense therapy, as probes for

CC diagnostic assays and for functional studies. BEER may be used to raise

CC antibodies and for identification of BEER modulators. BEER antagonists

CC may be used to increase bone mineral content for the treatment of

CC disorders such as osteopenia, osteoporosis, fractures and other disorders

XX associated with low mineral content.

XX

XX Sequence 213 AA;

XX

Query Match	95.2%;	Score 920.5;	DB 21;	Length 213;
Best Local Similarity	95.5%;	Pred. No. 7.2e-83;		
Matches 169; Conservative	2;	Mismatches 5;	Indels 1;	Gaps 1

QY 1 NDAETETIPELGEYPELPEPL -NNKTMNAENGSGPPHPELEIKDASEYSCRELHTRIV 35
Db 31 ndetellipeligeypepplenanktmrtaengagrpiphbpfetkdveysecrelhttrvt 90
QY 60 DGACRSRASPVTTELIVSGCGCPARLLLPNAIGRCKMWRSGDFEKCIPDRKRAORVOLLOCPG 119
Db 91 dgcrcsrkpytelivsgsgqgperlllpnaigrykwmpspdfcrlpdyreayvllpcp 150
QY 120 GAAPRARKYRIVASCKCRRLRHNOSLTKPFGEAPRQPTGSKLPPRARGKASRA 176
Db 151 gaaprarkyriyascckcrlrltrfmngslkdfpgeaarpqktrkprprargakanaga 207


```

1  APPLICANT: Watson, James D.
2  APPLICANT: Strachan, Lorna
3  APPLICANT: Sleeman, Matthew
4  APPLICANT: Orust, Rene
5  APPLICANT: Murlson, James Greg
6  TITLE OF INVENTION: Compositions Isolated From
7  TITLE OF INVENTION: and Methods for Their Use
8  FILE REFERENCE: 11000.101c1
9  CURRENT APPLICATION NUMBER: US/09/188,930A
10 CURRENT FILING DATE: 1998-11-09
11 NUMBER OF SEQ ID NOS: 348
12 SOFTWARE: FASTSEQ for Windows Version 3.0
13 SEQ ID NO 286
14 LENGTH: 206
15 TYPE: PR1
16 ORGANISM: Mouse
17 US-09-188-930-286

```

Query Match	34.7%;	Score 336;	DB 4;	Length 206;
Best Local Similarity	41.8%;	Pred. NO. 5.7e-28;		
Matches 76;	Conservative 29;	Mismatches 65;	Indels 12;	Gaps 6

[illegible]

RESULT 3
 US-08-468-847B-2
 : Sequence 2, Application US/08468847B
 : Patent No. 5780263
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Haslings, Gregg A. and Adams, Mark D.
 : TITLE OF INVENTION: Human CCR-like Growth Factor
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESS:
 :

1
2 CITY: ROSELAND
3 STATE: NEW JERSEY
4 COUNTRY: USA
5 ZIP: 07068
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: 3.5 INCH DISKETTE
9 COMPUTER: IBM PS/2
10 OPERATING SYSTEM: MS-DOS
11 SOFTWARE: WORD PERFECT 5.1
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/468,847B
14 FILING DATE: 6 June 1995
15 CLASSIFICATION: 435
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER:
18 FILING DATE:
19 ATTORNEY/AGENT INFORMATION:
20 NAME: MULLINS, J.G.
21 REGISTRATION NUMBER: 33, 073
22 REFERENCE/DOCKET NUMBER: 325800-444
23 TELECOMMUNICATION INFORMATION:
24

```

? TELEPHONE: 201-994-1700
?
? TELEFAX: 201-994-1744
?
? INFORMATION FOR SEC ID NO:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 206 AMINO ACIDS
?
? TYPE: AMINO ACID
?
? STRANDEDNESS:
?
? TOPOLOGY: LINEAR
?
? MOLECULE TYPE: PROTEIN
?
US-08-468-847B-2

```

Query Match	34.4%;	Score 333;	DB 1;	length 206;
Best Local Similarity	41.8%;	Pred. No. 1.2e-27;		
Matches	76;	Conservative	29;	Mismatches 65; Indels 12; Gaps 6;

[illegible]

RESULT 4
US-08-468-847B-20
; Sequence 20, Application US/08468847B

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 19:18:35 ; Search time 5549.44 Seconds
(without alignments)
3919.497 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301

Sequence: 1 agagcctgtctactgtgaag.....caatgaatcatgaccgaaag 2301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:51:52 ; Search time 438.73 Seconds
(without alignments)
918.818 Million cell updates/sec

Title: us-09-668-021-9

Perfect score: 642
Sequence: 1 atcgacgtccacgtgcct.....agctgagagacgtactag 642

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_0601:*

- 1: /SIDSL1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SIDSL1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SIDSL1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SIDSL1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SIDSL1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SIDSL1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SIDSL1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SIDSL1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SIDSL1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SIDSL1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SIDSL1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SIDSL1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SIDSL1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SIDSL1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SIDSL1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SIDSL1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SIDSL1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SIDSL1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SIDSL1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SIDSL1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SIDSL1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SIDSL1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	642	100.0	642	21	AAA29057
2	613.2	95.5	642	21	AAA94051
3	613.2	95.5	2301	21	Human DAN/Cerberus
4	613.2	95.5	2329	22	Human TGF-beta bin
5	611.6	95.3	2301	21	Human secreted pro
6	611.6	95.3	2301	21	Mutant human TGF-b
7	610	95.0	2301	21	Human TGF-beta bin
8	530	82.6	674	21	Human TGF-beta bin
9	503.2	78.4	638	21	Rat TGF-beta bindi
10	461.4	71.9	532	21	Murine TGF-beta bi
11	413.4	64.4	5680	21	Bovine TGF-beta bi
					Human DAN/Cerberus

12	413.4	64.4	9301	21	AAA29064	Human TGF-beta bin
13	399.6	62.2	1104	21	AAA94050	Human DAN/Cerberus
14	361	56.2	35628	20	AAA29063	Murine TGF-beta bi
15	142	22.1	198	20	AA440842	Human secreted pro
16	141.6	22.1	196	21	AAC05741	Human secreted pro
17	83.4	13.0	954	22	AAC99782	Skin cell cDNA, SE
18	83.4	13.0	962	21	AAZ61645	CDNA encoding murt
19	83.4	13.0	962	21	AAZ61736	CDNA encoding murt
20	83.4	13.0	962	22	AAC99578	CDNA encoding murt
21	83.4	13.0	962	22	AAC99669	Skin cell cDNA, SE
22	71	11.1	900	18	AAT47661	CDNA encoding smal
23	71	11.1	900	19	AAV37735	Human small CCN-11
24	71	11.1	1692	21	AAC79467	CDNA sequence of h
25	71	11.1	1710	20	AA484238	DNA encoding human
26	71	11.1	1719	21	AA404598	Human adult retina
27	71	11.1	1774	20	AA405771	Human sapiens fetal
28	57.8	9.0	412	22	AA666218	Novel human polynu
29	54	8.4	37856	21	AA666218	S. cellulosum DNA
30	52.2	8.1	2943	17	AAT16480	SAB virus gB glyco
31	48.2	7.5	114955	20	AA453491	Human adenosine A1
32	47.6	7.4	114955	20	AA453491	Human adenosine A1
33	47.2	7.4	71889	21	AA428349	Soraniangium cellulos
34	45.4	7.1	23666	12	AAQ10190	Cephalosporin anti
35	44.2	6.9	835	18	AAT77280	Lymantria dispar n
36	44.2	6.9	2791	16	AAO99776	Phosphomycin biosy
37	44.2	6.9	8867	21	AAA92058	Human Lhx3 genomic
38	43.6	6.8	5692	22	AAF32248	Streptomyces sp. C
39	43.4	6.8	1125	22	AAC84259	Signal transductio
40	43	6.7	4020	18	AAT91361	Orf virus genomic
41	42.6	6.6	4020	18	AAT91361	Orf virus genomic
42	42.6	6.6	5617	19	AAV57163	Partial human Notc
43	42.6	6.6	8091	19	AAV57001	Human Notch3 cDNA
44	42.6	6.6	43280	18	AAT80413	Tylosine synthase
45	42	6.5	117213	19	AA62176	HSV-2 strain SB5 C

ALIGNMENTS

RESULT 1	
ID AAA29057	standard. cDNA: 642 BP.
AC AAA29057;	
XX	
DT 12-SEP-2000	(first entry)
XX	
DE Vervet TGF-beta binding protein (BEER) cDNA.	
XX	
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KM BEER; gene therapy; antisense therapy; fracture; bone mineralization; ss.	
XX	
OS Cercopithecus pygerythrus.	
XX	
FH key	Location/Qualifiers
FT CDS	1..642
FT	/*tag= a
XX	/product= TGF-beta_binding_protein
PN	MO200032773-A1.
PD	08-JUN-2000.
XX	
PF 24-NOV-1999;	99WO-US27990.
XX	
PR 27-NOV-1998;	98US-0110283.
XX	
PA (DARW-) DARWIN DISCOVERY LTD.	
XX	
PI Brunow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
PI Van Ness J, Winkler DG;	
XX	
DR WPI; 2000-412321/35.	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 19:20:30 ; Search time 9526.57 Seconds
(without alignments)
3736.003 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301
Sequence: 1 agagcctgtcactcgaag.....caatgaatcatgcaccgaag 2301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Checked: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_om:*
9: gb_ov:*
10: gb_pat1:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_inv:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rsd:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2301	100.0	2323	89	AF326739 Homo sapi
2	2280.6	99.1	2329	9	AX056687 Sequence
3	2273.2	98.8	2296	89	AF331844 Homo sapi
4	2047.8	89.0	21501	89	AF326736 Homo sapi
5	2047.8	89.0	94752	85	AC003098 Homo sapi
6	2035.8	88.5	151780	72	AC055813 Homo sapi
7	1066.4	46.3	177744	75	AC073954 Homo sapi
8	613.2	26.6	642	89	AF326742 Cercopit


```

QY 781 attcatgttaaaltgctgcaaccagggagctgagacctccagcccttgaga 840
Db 781 ATTTCATTGTAATAGCTGCAACCCAGGGAGGGGGCTGAGACCTTCAGGCCCTGAGGA 840
QY 841 atcccgagcccgagagagcccccctcaagcccgcaagctgaggggtcccaagggcgagg 900
Db 841 ATCCCGAGCCCGAGAGAGCCCCCTCAAGCCCGCAAGCTGAGGGGTCCCAAGGGCGAGG 900
QY 901 gagggaaattgagagtcacagacatgagccagcccgccctctggggccgctacct 960
Db 901 GAGGGAATGAGAGTCACAGACACTGAGCCAGCACCCCGCTCTGAGGCCGCTACT 960
QY 961 ctgctgtctccacttcagagagagaaatgaaagacatttccagcccgctggggtttta 1020
Db 961 TTGCTGTCTCCACTTCAGAGAGGAGGAAATGGAAGCATTTTCACCGCCCTGGGGTTTAA 1020
QY 1021 agggagcggtgtgaggagtgaggaagtcagggactggttaagaagttgataagattcc 1080
Db 1021 AGGGAGCGGTTGTGGAGTGGGAAAGTCAGGAGACTGGTTAAAGATTGAGTAAGATTCC 1080
QY 1081 cccctgcaacctgcgtgcccacacagaaagcctgagggctgcccagagcaagagctggagg 1140
Db 1081 CCTTGCACGTCGCTGCCCCATCAGAAAGCCCTGAGGGGTGCCAGAGCAGCAAGACGTGGGG 1140
QY 1141 caactgtagatgtgtgttctagctcgtgctcgtccacttaactgtgtgttaacttgaa 1200
Db 1141 CAACGTGATGATGTGTGTTAGTCGTGCTGTGCCACTAACTTGTGTGTAACCTTTGAAAC 1200
QY 1201 tacacaattctcctcctcgagaccctcaattccacttgtlaaagtggaggggtgagggaa 1260
Db 1201 TACACAATTCCTCTTGGGAGCTCAATTTCCACTTTGTAAATGAGGGGTGAGGGTGGGAA 1260
QY 1261 tagatctcaagagagactatgcatatattccaaagactccaaatgcttttaaggg 1320
Db 1261 TAGATCTCGAGAGACTATGTCATATGATGATTCAGAGGACTCCAGTGCCTTTGATAGGG 1320
QY 1321 cagaggttgagagagagagagagagagagagagagagagagagagagagagagagag 1380
Db 1321 CAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 caaaggtcaactccagaaattcagagttgtatgtctctcttcaagagcaagaatgaa 1440
Db 1381 CAAGGTCACTCCAGAAATTCAGAGTTGTATGTCTCTCTTCTACAGCCCAAGAGTCAAAA 1440
QY 1441 caaagagagagagagagagagagagagagagagagagagagagagagagagagag 1500
Db 1441 CAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 tccctggaagagctatgtctctccagcctgctcccgagatgtttcgtctactccac 1560
Db 1501 TCCTTGAAGAGAGCTATGTCTCTCCAGCCTGCTTCCCGAGATGTTTGGTACTCTCAC 1560
QY 1561 cccctcatctcaagaaatacatcatcttggggtagaagagagagagagagagagagag 1620
Db 1561 CCTTCATCTCAAGAAAGAAATACATCATCTGAGGGGTAGAAAAGAGAGAGAGAGAGAG 1620
QY 1621 ggtggagagagtagaagatcacatccgcccgaacttcccaagagagagagagagagag 1680
Db 1621 GGTGGAGAGATGAGAAATACATCATCTGAGGGGTAGAAAAGAGAGAGAGAGAGAGAG 1680
QY 1681 acccatagcagatgtttaaagtcaacttccgagagagagagagagagagagagagag 1740
Db 1681 ACCCATAGCAGATGTTTAAAGTCACTTCGAGAGAGAAATGAAAGGTTTCAAGAGAGAG 1740
QY 1741 ccttgagagcccgagagagagagagagagagagagagagagagagagagagagagag 1800
Db 1741 CCTTGCAGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 cagcgcccttgcagagagagagagagagagagagagagagagagagagagagagagag 1860
Db 1801 CAGCGCCCTTGTGCCCCACACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860

```

```

QY 1861 tcttacatgtgatgcatatcttacactaaagaataltattggygggaaanaactacaagt 1920
Db 1861 TCTTACATGTGATGGCATATCTTACACTTAAAGAAATATTATTGGGGGAAAAATACAAAGT 1920
QY 1921 gctgtacatagtctggaagaactgcagagacataatagctgcagccaccaaactctttgaa 1980
Db 1921 GCTGTACATATGCTGGAAGAACTGCAGAGCATATAGCTGCCAACCCAAAATCTTTTGA 1980
QY 1981 aatcatttccagaacaacttacttctctgtgtagttttaattgttaaaaaaaaaaagt 2040
Db 1981 AATCATTTCAGAACACCTCTTACTTGTGTAGTTTATTTATTTTAAAAAAAAGT 2040
QY 2041 tttaagaagaagacatagacataatgaagccctgcagagactgctgttcttttgcaatc 2100
Db 2041 TTTAAGAGAGACATGACATATGAAAGCCGACGAGACTGTCGTTTTTTTGGCAATTC 2100
QY 2101 ttccagctggagactgttccacaagaatgagtagtgtttttaagaagttaagtaacat 2160
Db 2101 TTCCAGCTGGAGACTGTGCCAAGAAATGAAAGTAGTGTTTTAAAGATTAAGTATCAT 2160
QY 2161 attatttctcacttaagttattatgcaaaagtttctctgtagagatgacaatgct 2220
Db 2161 ATTATTTCTCACTTAACTTATTTATTCAAAAGTTTCTTTAGAGAAATGACAAATGT 2220
QY 2221 aatatgcttataatgaatgaagctgtctctccagagatccagagagaaatgtaataaag 2280
Db 2221 AATATGCTTATGAAATTAACAGCTGTGTTCTTCAGAGTCCAGAGACATNTTATTAAG 2280
QY 2281 acaatgaatcagagcgaag 2301
Db 2281 ACAATGAATCATGACCGAAG 2301

RESULT 2
AX056687 2329 bp DNA PAT 17-JAN-2001
LOCUS AX056687 Sequence 19 from Patent W00075317.
DEFINITION AX056687
ACCESSION AX056687
VERSION AX056687.1 GI:12309667
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2329)
AUTHORS Bolstein,D.A., Goddard,A., Gurney,A.L., Smith,V., Watanabe,C.K. and
WOOD,W.I.
TITLE Compositions and methods for the treatment of tumor
JOURNAL Patent: WO 0075317-A 19 14-DEC-2000;
Genentech, Inc. (US)
FEATURES
location/Qualifiers
source 1..2329
/organism="Homo sapiens"
BASE COUNT 587 a 645 c 612 g 485 t
ORIGIN
Query Match 99.1%; Score 2280.6; DB 9; Length 2329;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 14 ctggaaggtgagtgagcctctctctgtggtacatagagctccacatggccgtgtct 73
Db 28 CTGGAAGGTGAGTGCCTCTCTCTGTGTGTACAGACGCTCCCACTGGCCGTGTGCT 87
QY 74 cgtcgccctgtgttaacacagaccttccgtgtagtggagggccaggggtgtgagcgct 133
Db 88 CGTCGCTGTGTGTACACACAGCTTCGTTAGTGTGAGGGCCAGGGGTCGACGCGT 147
QY 134 caagatattccacagaaatacccgagctcgagagagtagcccgagctccacagga 193
Db 148 CAGATATATCCACGGAATCATCCCGAGCTCGAGAGTAGTACCCCGAGCTCCACCGGA 207

```

QY 194 gctggaacaacaagacatgaaccgagcgaggaacgagcgagcctccaccacc 253
|||||
Db 208 GGTGAGAACAAACAGACATGAACCGGGCGAGACGAGGGGGCTCCACACACC 267
QY 254 ctttagaaccaagcgctgctcagtagaagctgcgagctgacatccagctacgt 313
|||||
Db 268 ctttagaaccaagcgctgctcagtagaagctgcgagctgacatccagctacgt 327
QY 314 gacagaatggcgctgcgacgacgacgacgacgacgacgacgacgacgacgacg 373
|||||
Db 328 GACCGATGGCG 387
QY 374 cggccggcgcgctgcctgcgacgacgacgacgacgacgacgacgacgacgacg 433
|||||
Db 388 CG 447
QY 434 gcccgaattccgctgcacatcccgacgacgacgacgacgacgacgacgacgacg 493
|||||
Db 448 GCCCGACTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 507
494 cgttggatggagcg 553
|||||
508 CGGTGGTGAAGCG 567
QY 554 cctcaacgcgcttccacaacagtcggaagtcgaagactcgagacgagcgctcgcc 613
|||||
Db 568 CCTACACCGCTTCCACAAACATCGAGCTCAAGAGCTTGCGGACCGACCGCTGGCG 627
QY 614 gcaagaagcgcgagcg 673
|||||
Db 628 GCAGAGGGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 687
QY 674 gggagaacgcctactagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 733
|||||
Db 688 GGGAGAGCGCTCTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 747
QY 734 aaccgcgcccacattctcctcctcgcgcgctgctgctgctgctgctgctgctg 793
|||||
Db 748 ACCCGCGCCCGACATTTCTGCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 807
QY 794 tgcctgcaacccgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 853
|||||
Db 808 TGGCTGCAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
QY 854 gcaagggccccctcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 913
|||||
Db 868 GCAAGGCGCCCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 927
QY 914 gtcacagacactgaagcagcagcagcagcagcagcagcagcagcagcagcagcag 973
|||||
928 GTACAGACACTGAGCGACGCGCGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTG 987
QY 974 ttcagagggagcagaatctggaagcaatttcacgcgcctcggggttttaaggagcg 1033
|||||
Db 988 TTACAGAGGAGCAAGAAATGGAAGCATTTTACCGCCCTGGGGGTTTAAAGGAGCG 1047
QY 1034 ggaatggaagaatccagggagctgctgaagaagctggaalaattcccccctgcaact 1093
|||||
Db 1048 GGAATGGGAAGTCCAGGAGACTGTGTAAGAAAGTTGGAATGATCCCTTCACACTCG 1107
QY 1094 ctgcacataagaagcctgagcgctgctgcacagacacaagacttggggcaacttgaat 1153
|||||
Db 1108 CTGCGCATGAGAAAGCTGAGGCGTGCAGAGCAAGAACTGGGGCAACTGTATGATGT 1167
QY 1154 ggtcttaagctcctgcctgcacactaacttgcctgctgaacttgaactaacaactctcc 1213
|||||
Db 1168 GGTTCCTAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1227
QY 1214 ttggagacactcaattccaacttgaataataggtggaagtggaataggaactcgaag 1273
|||||
Db 1228 TTGGGAGCTCAATTTCCACTTTGTATAATGAGGTGGAGGTGGAGATGAGATCTCGAG 1287

QY 1274 agactatggcatalgatgtccaaaggaactccagtgctccttctgaatggcgagagtgagaga 1333
|||||
Db 1288 AGACTATTTGGCATATGATTCACAGAGACTCCAGTGCCTTTTGAATGGGAGAGGTGAGAGA 1347
QY 1334 gagaagagaagaagaagaagaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1393
|||||
Db 1348 GAG 1407
QY 1394 agaattcaagatgtgatgtctctcttcagacgacaaagatgaaanaaacaagaanaaa 1453
|||||
Db 1408 AGAATTCAGAGTGTGATGCTCTCTTCAGACGCAAGATGAAAGAAACAAACAGAAAGAA 1467
QY 1454 aaaaagtaagatctcatatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1513
|||||
Db 1468 AAAAGTAAAGAGCTATTTATGCTGACATTTTACGGCTGACAAACTCTGGAAGAAAC 1527
QY 1514 tatgtcgtctccagcgtgctcctccgagatglttggctactccaccctccatctcaa 1573
|||||
Db 1528 TATGCTGCTTCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1587
QY 1574 agaaataacatccatccatctgagtagaagaagagagagagagagagagagagagagata 1633
|||||
Db 1588 AGAATTAACATCATCATTTGGGGTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1647
QY 1634 gaaatacatccgcccccaacttccaaagagcagcatccctcccgagcccatagccatg 1693
|||||
Db 1648 GAATACATCCGCGCCCAACTTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1707
QY 1694 ttttaagtcacacttccgag 1753
|||||
Db 1708 TTTTAAAGTACACTTCCGAG 1767
QY 1754 agggagagacacatccaaactcagagagagagagagagagagagagagagagagagag 1813
|||||
Db 1768 AGGAGACAGCATCACAACACTCAGACACAGACATCCCTTTGAGACACCGCTTCTGC 1827
QY 1814 ccaacacacaggaacatctctgcctgaagaacagctcttactcctcttacaatgtag 1873
|||||
Db 1828 CCACCACTCAGGAGACATTTCTGCTGAGAAACAGCTTCTACTGCTTACATGTGAT 1887
QY 1874 ggcatalcttacaacaaagaatatattctgggggaaacacacacagctgctgataatgc 1933
|||||
Db 1888 GGCATATCTTACACTTAAAGATATATTGCGGGAAGAAACTACAGAGCTGATCATATGC 1947
QY 1934 tgaagaacatgcagagacataatagctgcacacacacacacacacacacacacacac 1993
|||||
Db 1948 TGAAGAACTGCAAGCATATATGCTGCGCCCAAAATCTTTTGAAGAAATCATTTCCAGA 2007
QY 1994 caacccttacttctcgtgtagtlttlaatgttaaaaaaaaglttlaaacagaagc 2053
|||||
Db 2008 CAACCTTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 2067
QY 2054 acatgacatataaagcctgcagagcgtgctgttcttggcaatcttccacgctgagac 2113
|||||
Db 2068 ACATGACATATGAAGCCCTGCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2127
QY 2114 ttgtccacaagaatgaagaatgagttttaaagaagatgaatgaatgaatgaatgaatga 2173
|||||
Db 2128 TTGTCCACAAGATGAAGATGAGTGTTTTAAAGAGTTAGATTAATTTATTTCTCA 2187
QY 2174 cttgaattatatacaaaagtttctctfagagaaatgacatgtaataatctgctctat 2233
|||||
Db 2188 CTTAAAGTTATTAATGAAGATTTTCTTGTAGAGATGCAATGTTAATATTGCTTAT 2247
QY 2234 gaattacagctgtctctccagagtcagagacatgttataaagaacatgaacatgacatg 2293
|||||
Db 2248 GAATTAACAGTCTGCTTCCAGAGTCCAGAGATGTTAATTAAGACATGAATCATG 2307
QY 2294 accgaaa 2300
|||||
Db 2308 AAAAAA 2314

```

RESULT 3
AF331844 2296 bp mRNA PRI 06-MAR-2001
LOCUS Homo sapiens SOST (SOST) mRNA, complete cds.
DEFINITION AF331844
ACCESSION AF331844.1 GI:13236417
VERSION 1
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2296)
AUTHORS Balemans,W., Ebeling,M., Patel,N., van Hul,E., Olson,P.,
Diozegui,M., Lacca,C., Wuyts,W., van den Ende,J., Williams,P.,
Paez-Alves,A.F., Hill,S., Bueno,M., Ramos,F.J., Tacconi,P.,
Dijkers,F.G., Stratakis,C., Lindpaintner,K., Vickery,B.,
Foerzler,D. and Van Hul,W.
Increased bone density in sclerosteosis is due to the deficiency of
a novel secreted protein (SOST)
Hum. Mol. Genet. 10 (5), 537-543 (2001)
1181578
JOURNAL PubMed
REFERENCE 2 (bases 1 to 2296)
AUTHORS Balemans,W., Ebeling,M., Patel,N., Vickery,B., Foerzler,D. and Van
Hul,W.
Direct Submission
TITLE Submitted (22-DEC-2000) Medical Genetics, University of Antwerp,
JOURNAL Universiteitsplein 1, Antwerp B2610, Belgium
FEATURES
SOURCE location/Qualifiers
1..2296
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q12-q21"
1..2296
/gene="SOST"
38..679
/gene="SOST"
/gene="SOST"
/note="Mutations may cause sclerosteosis"
/codon_start=1
/product="SOST"
/protein_id="AAK16158.1"
/db_xref="GI:13236418"
/translation="MOLPLALCIYGLIVHTAFRVYEGOGWAKNDATETIIPLEGXP
EPPELENNKTKMRAENGSRPHHPFEKDYSEI SCHELFTRIVYDGRPRKAPYTE
LVSCGCGGARLLPNAIGRGKMWPRSGDFRCTDRIRARVOLICGGGAPKRVK
LVASCKCKRLTRFHNQSELDKDFGEAARPDQGRKRPBRASANAQALEENAY"
CDS
gene
CDS
ORF COUNT 576 a 631 c 607 g 482 t
Query Match 98.8%; Score 2273.2; DB 89; Length 2296;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2286; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

OY 311 cgtacccgatggccgtgctgcagcagcccaagccggtcaccagctgtgtctccgccca 370
DB 301 cgtacccgatggccgtgctgcagcagcccaagccggtcaccagctgtgtctccgccca 360
OY 371 gtcgagcccgagcgcctgtctgcacaaagccatcgccgagcagctgtgtgcagctag 430
DB 361 gtcgagcccgagcgcctgtctgcacaaagccatcgccgagcagctgtgtgcagctag 420
OY 431 tgggcccgaacttcgctgcacatcccccagccgctacccgagcagcgtgcagcgtgtg 490
DB 421 tgggcccgaacttcgctgcacatcccccagccgctacccgagcagcgtgcagcgtgtg 480
OY 491 tcccggtgtgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 550
DB 481 tcccggtgtgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
OY 551 gcgcctcccgcttcgcacaaacagctcgagctcaagagcttcggagacccgagcgcctcg 610
DB 541 gcgcctcccgcttcgcacaaacagctcgagctcaagagcttcggagacccgagcgcctcg 600
OY 611 gccgcaagaaggccggaagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 670
DB 601 gccgcaagaaggccggaagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 660
OY 671 gctggagagagccttaagagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 730
DB 661 gctggagagagccttaagagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
OY 731 tgaacccgcgcgcacacattctgtctctctgtgcgcgtgtgtgtgtgtgtgtgtgtgt 790
DB 721 tgaacccgcgcgcacacattctgtctctctgtgcgcgtgtgtgtgtgtgtgtgtgtgt 780
OY 791 aaatgcttcgaacccgaagcagaggggtgtgtgaccttcgaagccttcgaagaaatcccg 850
DB 781 aaatgcttcgaacccgaagcagaggggtgtgtgaccttcgaagccttcgaagaaatcccg 840
OY 851 ccggaagggcccccctcagcccgccagctgaagggttcaccaagggcagggaggaattg 910
DB 841 ccggaagggcccccctcagcccgccagctgaagggttcaccaagggcagggaggaattg 900
OY 911 aagagcaagaacacttgagcagcagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 970
DB 901 aagagcaagaacacttgagcagcagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
OY 971 caattcagaaggcaggaatggaacatttcacgcgccttgagggttttaaggagagcgt 1030
DB 961 caattcagaaggcaggaatggaacatttcacgcgccttgagggttttaaggagagcgt 1020
OY 1031 gtggaggtgggaaggtccagagactgtttaagaagtttgataagaatcccccctgcacc 1090
DB 1021 gtggaggtgggaaggtccagagactgtttaagaagtttgataagaatcccccctgcacc 1080
OY 1091 tgcgtcccatcagaagaagcctgaagcgtgccagagcaacaagactgagggcaactgtaga 1150
DB 1081 tgcgtcccatcagaagaagcctgaagcgtgccagagcaacaagactgagggcaactgtaga 1140
OY 1151 tgggtttctatgcttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 1210
DB 1141 tgggtttctatgcttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 1200
OY 1211 tcccttcgagcctcaatttcacatttgaatgaaggtgaggtgaggtgaggtgaggtgag 1270
DB 1201 tcccttcgagcctcaatttcacatttgaatgaaggtgaggtgaggtgaggtgaggtgag 1260
OY 1271 aggaactatgtgcatatgattcagaagactcaagtgcccttgaatggagaggttag 1330
DB 1261 aggaactatgtgcatatgattcagaagactcaagtgcccttgaatggagaggttag 1320
OY 1331 agagaagagagaaagagagagaaatgaatgaatgaatgaatgaatgaatgaatgaatga 1390
DB 1321 agagaagagagaaagagagagaaatgaatgaatgaatgaatgaatgaatgaatgaatga 1380

```

QY	1391	tcgaagatccagaagctgtagtgcctctctctgaagccaaagatgaaacaaacaga	1430
Db	1381	TCGAGATTCCAGATGGTTGATGAGTCTCTCTCTGACACGGCAAGATGAAAAACAACGAAA	1440
QY	1451	aaaaaagaataaagctatctattatggcctgacatcttcggctgacaaactccctgga	1510
Db	1441	AAAAAAGTAAAGAGTCTATTATATGCGTACATATTTACGGCTGACAACTCTGGAGA	1500
QY	1511	agctatgcgctcccgagcctggctcccgagatcttggtgactccacacctcaact	1570
Db	1501	AGCTATGCGCTTCCACACTGGCTTCCCGGATTTTGGCTACCTCCACCTCTCATCT	1560
QY	1571	caagaataacatcatcatcttgaggtagaaaaagaagagctccgaagtgctgagag	1630
Db	1561	CAAGAATAATATCATCTCATTGGGCTAAGAAAAAGAGAGGCTCCGAGGCTGGGGAGG	1620
QY	1631	atgaatcatcatccggccccaacttccaaagaagacatccctcccgacccatagcc	1690
Db	1621	ATGAAATCATCATCGGCCCACTCCCAAAAGAGAGATCCCTCCCGCACTCATACC	1680
QY	1691	atgttttaagtcaactctccgaagaagtgaaagtgctcaagaactgcttgagc	1750
Db	1681	ATGTTTAAAGTCACTCTCCGAAGAATAATGAAAGTTCAAGGACATGGCTTCCAGCC	1740
QY	1751	ccgaggagagcgcacatcaaaactccagaccagaacatccctctgagacccgcttc	1810
Db	1741	CCGAGGAGACACACCTCATCAAACTCACAGACAGCATCCCTTTGAGACACCGCTTC	1800
QY	1811	tggcccaacactccagggacacttctgctctgaaaaagcctctctctgcttcaagt	1870
Db	1801	TGCCACACACTACGGACACATTTCTGCTTAAAMACAGCTTCTTACTGCTCTACAGT	1860
QY	1871	gatgcatatccttcaacataaagaatattatgggggaaaaactcaagtgctgata	1930
Db	1861	GATGCAATATCTTACCTAAAGAAATATTATGGGGGAAAAACTACAGTCTGATACATA	1920
QY	1931	tgtctgaaactgcagagacataatagctgcacccaaaaactctttgaaatcttcc	1990
Db	1921	TGCTGAGAAACGACAGCATTAATA-CTGCCACCCAAAATCTTTTGAAATCATTTCC	1979
QY	1991	agacaaactcttactctctcgctgtagtctttaaagtgttaaaaaaaagtttaacsa	2050
Db	1980	AGCAACCTCTTACTCTCTGTGATGTTTTTAATTTTAAAAAAAAGTTTTTAACAGA	2039
QY	2051	agcacatgacatcagaaagcctgcagagactgctgttttttgcaactctccacgtg	2110
Db	2040	AGCACATGCATTAAGAAAGCCTGCAGCATGCTGCTTTTTTGGCAATCTTCCAGTGG	2099
QY	2111	gactgtccacaagaatagaagtaggtttttaagaagttaagttacatatcttctc	2170
QY	2100	GACTGTCCACACAAAGAAAGTAGTGTTTTTAAAGAGTTAAGTTCATATTTATTTTC	2159
QY	2171	tcacttaagtattatcagaaagtttctctgtgtagaaatgacaaagttaactgctt	2230
Db	2160	TCACTTAAGTTATTATGCAAAAGTTTTCTGTGTGAGAAATGCAATGTTAATTTGCTT	2219
QY	2231	tatgaaatcaacgctctctcccaagatccagagacatggttaataaagaacatgaac	2290
Db	2220	TATGAATTAACAGTCTGTCTTCCAGAGTCCAGAGACATTGTTAATAAGACAAATGATC	2279
QY	2291	atgaccagaa 2300	
Db	2280	ATGCAAAAAA 2289	

RESULT	4
AF36736	
LOCUS	AF36736 21501 bp DNA PRI
DEFINITION	Homo sapiens sclerostin gene, complete cds.
ACCESSION	AF36736
VERSION	AF36736.1 GI:13161010
KEYWORDS	.
SOURCE	human.

[illegible]

QY	639	cgcccccgagcgcccaaaacccaaacgaagcgagcttgagaagccttactagaagcccgccc	698
Db	15310	CGCGCCCGGAGCCCAAGGCCAACAGSCCAGCTGTGAGAACGCTACTTGAACCCGCCC	15369
QY	699	gcgcacctcccaacgagcgagcgcccgagccctgaaccgagcccaatctctctctc	758
Db	15370	GCGCCCTCCCAACGCGCGGGCCCCCGCCTTAACCGCGGCCACATTTCTGTCCTC	15429
QY	759	tgcgcgtggtttgatgttattatcttacttgtaagtccctgcgaaccgaagcgagtggtc	818
Db	15430	TGCGCGGGTTTTGATTGTTTTATTATTCTAGTGAATGCTGTCAACCCAGGCGAGGGGCT	15489
QY	819	gaagacctccagggcccttgagaaatcccgagcgcccggaagggccctcagcccgcaac	878
Db	15490	GAGAGCTCTCAGAGCCCTGTAGGAAATCCCGGGCGCGGCAAGGCCCTCTACGCCGCAAC	15549
QY	879	tgaaggggtcccaagggcgagggagggaaattgaagatcaacagacatgagccagcagcc	938
Db	15550	TGAGGGGTCCCAAGGGGCGAGGAGGAATGAGACGTACAGACACTGAGCCACGCAAGCC	15609
QY	939	ccgcctcttggggcgagcctacacctttgccttgctccaaactcgaaggagcgaaatgagaa	998
Db	15610	CCGCTCTGGGGCGGCTACCTTTGCTGTGCTCCACTTCAGAGGAGGCAAAATGSAAGCA	15669
QY	999	ttttcaaccgcccgtggggttttaaggaagcggtgtgagagtgaggaagtccaagatggt	1058
Db	15670	TTTTTACCGCCCTGGGGTTTTTAAGSAGGCGGTGGGAGTGGGAATCCAGGAGCTGGT	15729
QY	1059	taagaaagtltgaaataagatltccccccttgacactcgctgcctcatgaagaagcctgaagcgt	1118
Db	15730	TAAAGAAAGTTGGATGATGATTTCCCTTCACCTGTGCTCCATCAGAAACCTCGAGCGCT	15789
QY	1119	gcccagaagcaagaactctgggggcaactctagatgtggtttctatgctctgctgcacat	1178
Db	15790	GCCAGAGCAACAAGACTGGGGGCACACTGTAGATGTGGTTCTATCTCTGGCTCTGCCACT	15849
QY	1179	aactgtcgtgttaacttgaaatacaaatctctcttcggagactaaattccacttgt	1238
Db	15850	AACCTGTGTGTAACTTGAACCTTAACACAACTTCTCTCTGGGAACTCAATTTCCACTTGTG	15909
QY	1239	aaaaatgggttggaggttgggaatagatctcgaagagacatttgcatatgattccaaag	1298
Db	15910	AAAATGAGGGGTGAGAGGTGGGAATAGAACTCGAGAGAGCTTTGTGCATATGATTTCCAAAG	15969
QY	1299	actccagatgcttttbaatgggacagagtgtgagaagagagaagagaagagaatgaa	1358
Db	15970	ACTCAGATGCCCTTTTGAATGGGCAGAGTGAAGAGAGAGAGAAAGAGAGAGATGAA	16029
QY	1359	tgcagtgatgcatatcatcagltgcgaagttcaacttccagaatccagatgtgatgtctct	1418
Db	16030	TGCGATTTGCAATTTATTTCACTGCGCAAGGTCACTTCCAGAAATTCAGAGTTGTGATGCTCTC	16089
QY	1419	tctacacgcaaaagatgaaaaaaaaacaaacgaaaaaaaagaaagagttctattatggtc	1478
Db	16090	TCTGACAGCCAAAGATGAAAAAACAAGAAAAAAGAAAAAGTAAAGAGTCTATTATATGGCT	16149
QY	1479	gaacatttcacgctgaacaaactccttgaaaaaagctatgctgtctcccaacgctggttcc	1538
Db	16150	GACATATTTCGCGCTGACAACTCCTGGAAGAGTATGCTGCTTCCAGCCTGCGCTTCC	16209
QY	1539	ccggaatgttggtactctccaccctccatctcaagaataacatccatcttgggta	1598
Db	16210	CCGATATTGGCTACTCTCACCCCTCCATCTCAAAAGAAATATCATTTGGGGTA	16269
QY	1599	gaaagaagagaggttccgaaggttggttgagagataagaaatacacaatccgcccacatccc	1658
Db	16270	GAAGAAGAGAGGGTCCGAGGGTGTGTGGAGGATGAAATATACATCCGCCCAACTTCCC	16329
QY	1659	aaagaagagatccctcccccgaaccatagacatgttttaagtcacacttcgaagaaga	1718
Db	16330	AAAGAGAGATCCCTCTCCCGACCAATATGCAATGTTTTAAAGTCACTTCTCGAAGAGAA	16389
QY	1719	gtgaagaagttcaagaaactgctcttgacgcccgaagggagcagcatcacaactcaca	1778

Db	16390	GTGAAGGTTCAAGSACATCGCTTGACGGCCCGAGGGAGAGCAGCATACAAACCTCACA	16449
Db	1779	gaccagcaatccctttggagaaacgcgcttcgcccaacactaecggaacattctgc	1838
Db	16450	GACCACACATCCCTTTTGAAGACACCGCCCTTTCGCCACACTCAGCAGACATTTCTGC	16509
OY	1839	ctgaaacaagctcttacccttaccacgtaatgagcaatcttaccactlaaagaatat	1898
Db	16510	CTGAAACACAGCTCTCTACCTCGCTCTTCACATGATGCAATCTTACACTTAAAGAAATAT	16569
OY	1899	tattgggggaaaaactaaagtgtgtacatatgctggagaacttcagagcataatagct	1958
Db	16570	TATTGGGGGAAAAACTACAAATGCTGTACATATGCTGAGAAACTCAGACGATTAATAGCT	16629
OY	1959	gccaccacaaaactcttttgaataatcatcttcacagacacaccttaacttcgtgtagtt	2018
Db	16630	GCCACCCAAAATCTTTTGAAGATTCATTTCCAGACAACTCTTACTTTCGTGTAGATT	16689
OY	2019	ttaattgttaaaaaaaaaggttttaacagagaagcacaagacataagaaacgcgcagaga	2078
Db	16690	TTAAATGTTTAAAAAAAAGGTTTAAACAGAGCAGCATGACATVTGAAACCCGTCAGCA	16749
OY	2079	ctgtgcgttttttggcaattcttcacgcgtagacttgcacgaagtgaagttaagt	2138
Db	16750	CTGTCCTTTTTTTGGCAATCTTCCACGTGGAGCTTGTCCCAAGATVGAAGTAATGG	16809
OY	2139	ttttaaagatgaatgatacatattatcttctacttaagtattatgcaaaagtltt	2198
Db	16810	TTTTTAAAGGTTAAGTATACATATTATTATTTCTCACTTAAGTATTATTTGCAAAAGTTTT	16869
OY	2199	tctgttagagaatgacaatgtaatatgttcttataatgaacagtcgcttccagag	2258
Db	16870	TCTGTGAGAAATGACAAATGTTAATTTGCTTATTAATTAACAGTCTGTTCTTCAGAG	16929
OY	2259	tccagagacatgtaataaagaacaaagaaatcatgacgcgaagag	2301
Db	16930	TCCAGACACTTGTATTATTAAGACAAATGATCATGACCCAAAG	16972
RESULT	5		
LOCUS	AC003098	94752 bp	DNA PRI
DEFINITION	Homo sapiens chromosome 17, clone HRC905N1, complete sequence.		29-JAN-1998
ACCESSION	AC003098		
VERSION	AC003098.1	GI:2822155	
KEYWORDS	HTG		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 94752)		
	Barren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P. and		
	Lander,E.		
	Homo sapiens chromosome 17, clone HRC905N1		
	Unpublished		
TITLE	2 (bases 1 to 94752)		
JOURNAL	Barren,B., Fasman,K., McKernan,K., Munro,C., Nusbaum,C.,		
REFERENCE	Richardson,P., Lander,E., Baldwin,J., Barne,N., Cantu,C., Chang,A.		
AUTHORS	Cooke,P., Daly,M.J., Devon,K., Dewar,K., Dunette,B., Forrest,C.,		
	Gage,D., Gensheimer,S., Gerald,J.K., Gilmarlin,T., Hagos,B.,		
	Halphen,I., Harris,K., Howland,J.C., Huang,J., hai,L., Jacobot,J.		
	Kilby,A., Lane,M., Mackenzie,J., Margolis,N., McDermott,J.,		
	Mold,M., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A.		
	Peterson,K., Roberts,D., Rollins,G., Sarrak,A., Shu,P., Shyam,R.		
	Stillewell,J., Stone,C., Strickland,C., Sydney,K., Tang,L.,		
	Zemeeva,I. and Zody,M.		
	Direct Submission		
JOURNAL	Submitted (10-NOV-1997)		
REFERENCE	Research, 320-Centering Street, Cambridge, MA 02141, USA		
AUTHORS	3 (bases 1 to 94752)		
	Barren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P.,		
	Lander,E., Allen,N., Baker,D., Baldwin,D., Barne,N., Beckert,I., R.,		

Boutwell,C., Byrne,S., Cantu,C., Castle,A., Cooke,P., Daly,M.J.,
 Depayre,E., Devon,K., Dewar,K., Durette,B., Etanadi,S.,
 Ferreira,P., Forrest,C., Gage,D., Gardyna,S., Gensheimer,S.,
 Geraghty,K., Gilmartin,T., Gray,D., Haggos,B., Harris,K.,
 Horton,L., Howland,J.C., Hul.L., Jacotot,L., Linton,L.,
 Mackenzie,J., Marquis,N., McEwan,P., McGuck,A., Meldrim,J.,
 Molla,M., Morris,N., Morrow,J., Nachman,A., Naylor,J., O'Connor,T.,
 Pavlin,B., Peterson,K., Ranganath,S., Riley,R., Roberts,D.,
 Rollins,G., Rossello,R., Roy,A., Shyam,R., Soohoo,S., Stillwell,J.,
 Stone,C., Strickland,C., Sydney,K., Tang,L., Vassiliev,H., Vo,A.,
 Wagner,A., Wheeler,J., Wu,Y., Ye,W.J., Zemtseva,I., Zhao,J. and
 Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (29-JAN-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Jan 29, 1998 this sequence version replaced gi:2655504.
 The Staden databases, finishing information, and all
 chromatographic files used in the assembly of this clone are
 available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. &
 Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

FEATURES
 source
 Location/Qualifiers

```

1..94752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HRC905N1"
/clone_11b="Peter de Jong/ human PAC library"
/map="17"
/chromosome="17"
86..439
/rpt_family="LINE2"
complement(972..1262)
/rpt_family="Alusx"
complement(1884..1912)
/rpt_family="(CA)n"
1912..1939
/rpt_family="(GA)n"
2633..2681
/rpt_family="MIR"
2701..3002
/rpt_family="Alusx"
complement(3874..4008)
/rpt_family="LINE2"
4513..4543
/rpt_family="AT-rich"
complement(5202..5337)
/rpt_family="(GA)n"
complement(5310..5406)
/rpt_family="MIR"
complement(5829..5868)
/rpt_family="GC-rich"
6485..6592
/rpt_family="MIR"
complement(6973..7029)
/rpt_family="MIR"
complement(10402..10426)
/rpt_family="GC-rich"
complement(11421..11496)
/rpt_family="LINE"
complement(11505..11804)
/rpt_family="Alusx"
11811..11879
/rpt_family="AT-rich"
11880..12093
/rpt_family="Alusx"
complement(12104..12476)
/rpt_family="LINE2"
12478..12773
/rpt_family="Alusx"
complement(12775..12917)
/rpt_family="LINE2"

repeat_region complement(13088..13510)
repeat_region /rpt_family="MLTIC"
repeat_region complement(13544..13847)
repeat_region /rpt_family="Alusx"
repeat_region complement(14110..14776)
repeat_region /rpt_family="LINE1"
repeat_region complement(14777..15077)
repeat_region /rpt_family="Alusx"
repeat_region complement(15079..15282)
repeat_region /rpt_family="LINE1"
repeat_region 15954..16024
repeat_region /rpt_family="AT-rich"
repeat_region 18031..18178
repeat_region /rpt_family="MIR"
repeat_region 20398..20769
repeat_region /rpt_family="LINE"
repeat_region 20893..21190
repeat_region /rpt_family="Alusg"
repeat_region complement(21275..21473)
repeat_region /rpt_family="Alusg"
repeat_region complement(21485..21556)
repeat_region /rpt_family="(CA)n"
repeat_region 21567..21906
repeat_region /rpt_family="MER2"
repeat_region 21907..21952
repeat_region /rpt_family="MLT2CB"
repeat_region 21974..22115
repeat_region /rpt_family="MIR"
repeat_region 22238..22540
repeat_region /rpt_family="Alusx"
repeat_region 22577..22633
repeat_region /rpt_family="LINE2"
repeat_region complement(23190..23270)
repeat_region /rpt_family="LINE2"
repeat_region 24194..24324
repeat_region /rpt_family="FLAM_A"
repeat_region 24328..24643
repeat_region /rpt_family="Alusg"
repeat_region 25542..25698
repeat_region /rpt_family="LINE2"
repeat_region 25805..26001
repeat_region /rpt_family="MIR"
repeat_region 26060..26174
repeat_region /rpt_family="LINE2"
repeat_region 26343..26649
repeat_region /rpt_family="Alusx"
repeat_region complement(27491..27783)
repeat_region /rpt_family="Alusg"
repeat_region 27862..27906
repeat_region /rpt_family="LINE2"
repeat_region 28068..28114
repeat_region /rpt_family="MIR"
repeat_region complement(28329..28360)
repeat_region /rpt_family="AT-rich"
repeat_region 28371..28652
repeat_region /rpt_family="Alusg"
repeat_region complement(29498..29604)
repeat_region /rpt_family="GC-rich"
repeat_region 30074..30181
repeat_region /rpt_family="MIR"
repeat_region complement(30181..30249)
repeat_region /rpt_family="LINE2"
repeat_region 30444..30517
repeat_region /rpt_family="MIR"
repeat_region complement(31852..32135)
repeat_region /rpt_family="Alusg"
repeat_region complement(32675..33003)
repeat_region /rpt_family="MLT1A1"
repeat_region complement(33042..33082)
repeat_region /rpt_family="LINE2"
repeat_region 33363..33459
repeat_region /rpt_family="MER5B"
repeat_region complement(35475..35524)
```

```
repeat_region /rpl_family="MLT2D"
35495..35591
/rpl_family="(GA)n"
complement(35561..35673)
repeat_region /rpl_family="MLT2D"
35673..35935
complement(35678..35935)
repeat_region /rpl_family="MER66B"
35935..36233
complement(35936..36233)
repeat_region /rpl_family="Alusg"
36233..36449
complement(36236..36449)
repeat_region /rpl_family="MER66B"
36449..36867
complement(36450..36867)
repeat_region /rpl_family="MLT2D"
36867..37460
complement(36876..37460)
repeat_region /rpl_family="LIMD2"
37460..37777
complement(37477..37777)
repeat_region /rpl_family="Alusx"
37777..37948
complement(37804..37948)
repeat_region /rpl_family="L1"
37948..38130
/rpl_family="LIMD1"
repeat_region /rpl_family="LINE2"
38130..38253
/rpl_family="LINE2"

Query Match 89.0% Score 2047.8; DB 85; Length 94752;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2061; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 219 cgggcgagagacgagagcgagcctcccaacacccctttagaaccaagacgtgccag 278
DB 6340 CAGGGGGGCGCAGGGGGGCGCCACCTCAGCGGCGCTCTCTCCACAGAGGTGCCAG 6281
QY 279 tagagctgcgagagctgaactcccgctacgtgacagatggcggtgagcgagcgcc 338
DB 6280 TAAGCTGCGGAGCTGCACTTCAACCCCTACGTACGATGAGGCGGTGCGGAGGCGC 6221
QY 339 aagccggtacacgagctgtgtgtctccgagcagtcgagccgagcgagcgtctgtcccaac 398
DB 6220 AAGCGGTGACCGAGCTGTGTCTCCGCGCAGTCCGCGCGCGCGCTGCTGCCAAC 6161
QY 399 gcatcgccgagcgaagtgtgtgagcacttagtggccgagcttcgctgatacccgac 458
DB 6160 GCCATGCGGCGGCAAGTGTGTGACCTAGTGGAGCCCGACTTCGCTGATCCCGAC 6101
QY 459 cgttacccgagcagcgctgtcagctgtgtgtccggttgggtgagcgcgcgcgcgcc 518
DB 6100 CGCTACCGCGCGCAGCGCGCTGCTGTCTCCGCTGAGAGCGCGCGCGCGCGC 6041
QY 519 aaggtgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 578
DB 6040 AAGGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5981
QY 579 gaagctcaagagacttcggagacgagcggtcggcgagcagagagcgagagcgagcgcc 638
DB 5980 GAGCTCAAGAGACTTCGGGACCGAGGCGCGTGGCGCGAGAAAGGCGCGGAGCGCGCC 5921
QY 639 cggcgccgagagcgcaaaagcgaacgagcgagctgtgagagagcctactatagacgcgcc 698
DB 5920 CGGCGCGGAGCGCAAAACCAACGAGCGGTGAGAAAGCGCTACTATAGCGCGCGCC 5861
QY 699 ggcgcctcccaacgagcgagcgcccgagccttgaacccgagcgcccaattctgtcttc 758
DB 5860 GCGCGCTCCCAACGAGCGCGCGCGCGCGCGCTGAAACCGCGCGCGCATTTCTGTCTC 5801
QY 759 tgcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 818
DB 5800 TGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5741
QY 819 gtagaccttcagagccttcgagagatcccgagcgagcgagcgagcgagcgagcgagcg 878
DB 5740 GAGACCTTCAAGCGCTTCAGGAATCCCGGCGCGCGCAAGCGCGCGCGCGCGCGCAGC 5681
QY 879 tgaaggggtcccaacgagcgagcgagagagagagagagagagagagagagagagag 938
```


Db 4600 GCCACCCAAAAATCTTTTGAAATCATTTCCAGAACCTTCTACTTCTGTGATTT 4541

QY 2019 ttaattgtaaaaaaaagtttaaaagaagcaatgatcatgtgaagcgcgcagga 2078
 |||||||

Db 4540 ttaattgtaaaaaaaagtttaaaagaagcaatgatcatgtgaagcgcgcagga 4481
 |||||||

QY 2079 ctgtgctgttttttgcaattcttccacgtgagctgtgccagaagaagtaagtcg 2138
 |||||||

Db 4480 ctgtgctgttttttgcaattcttccacgtgagctgtgccagaagaagtaagtcg 4421
 |||||||

QY 2139 ttttaaaagttaagttacatattatttttccactaagttattttgcaaaagttt 2198
 |||||||

Db 4420 ttttttaaaagtttaagttacatattatttttccactaagttattttgcaaaagttt 4361
 |||||||

QY 2199 tctgtagaagatgaacatgttaattgtcttaagaaatcaacagtcgtcttccagag 2258
 |||||||

Db 4360 tctgtagaagatgaacatgttaattgtcttaagaaatcaacagtcgtcttccagag 4301
 |||||||

QY 2259 ttcagagacatgttaaaagaacatgatcatgaccgaaag 2301
 |||||||

Db 4300 ttcagagacatgttttaaaagaacatgatcatgaccgaaag 4258
 |||||||

AC055813 151780 bp DNA HTG 22-JUL-2000
 LOCUS Homo sapiens chromosome 17 clone RP11-209M4 map 17, WORKING DRAFT
 DEFINITION
 AC055813
 AC055813
 AC055813.3 GI:9369521
 HTG: HTGS_PHASE1, HTGS_DRAFT.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 151780)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 17, clone RP11-209M4
 Unpublished
 2 (bases 1 to 151780)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barne,N., Bastien,V., Beda,F.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campiano,A., Castle,A., Choquet,I., Colangelo,M., Collins,S.,
 Collamore,A., Cooke,P., Dextrellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,I.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Klandt,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Larocque,K., Lamazares,R., Landers,T., Lenoczky,J.,
 Levine,R., Lien,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McCurt,A., McKernan,K., McNeeters,R.,
 Meltrin,J., Meneses,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testaye-S., Theodore,J., Tirrell,A., Travers,M., Triggiani,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zakhoun,J., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT
 Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 22, 2000 this sequence version replaced gi:8920653.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
 Center project name: L7949
 Center clone name: 209.M.4
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 94% of reads
 Sequencing vector: Plasmid; n/a; 0.1% of reads
 6.15883306320908Chemistry: Dye-terminator Big Dye; 100% of
 reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 139835 bases at least Q40
 Consensus quality: 145751 bases at least Q30
 Consensus quality: 148014 bases at least Q20
 Insert size: 155000; agarose-fp
 Insert size: 149480; sum-of-ctgifs
 Quality coverage: 3.9 in Q20 bases; agarose-fp
 Quality cov.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 24 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 1283: contig of 1283 bp in length
 * 1284 1383: gap of 100 bp
 * 1384 2931: contig of 1548 bp in length
 * 2932 3031: gap of 100 bp
 * 3032 4160: contig of 1129 bp in length
 * 4161 4260: gap of 100 bp
 * 4261 6377: contig of 2117 bp in length
 * 6378 6477: gap of 100 bp
 * 6478 8584: contig of 2107 bp in length
 * 8585 8684: gap of 100 bp
 * 8685 10624: contig of 2140 bp in length
 * 10625 10924: gap of 100 bp
 * 10925 13710: contig of 2786 bp in length
 * 13711 13810: gap of 100 bp
 * 13811 17466: contig of 3656 bp in length
 * 17467 17566: gap of 100 bp
 * 17567 19747: contig of 2181 bp in length
 * 19748 19847: gap of 100 bp
 * 19848 22997: contig of 3150 bp in length
 * 22998 23097: gap of 100 bp
 * 23098 26840: contig of 3743 bp in length
 * 26841 26940: gap of 100 bp
 * 26941 32331: contig of 5391 bp in length
 * 32332 32431: gap of 100 bp
 * 32432 37589: contig of 5158 bp in length
 * 37590 37689: gap of 100 bp
 * 37690 43814: contig of 6125 bp in length
 * 43815 43914: gap of 100 bp
 * 43915 49593: contig of 5681 bp in length
 * 49594 49693: gap of 100 bp
 * 49694 56982: contig of 7287 bp in length
 * 56983 57082: gap of 100 bp
 * 57083 63903: contig of 6821 bp in length
 * 63904 64003: gap of 100 bp
 * 64004 73106: contig of 9103 bp in length
 * 73107 73206: gap of 100 bp
 * 73207 81472: contig of 8266 bp in length
 * 81473 81572: gap of 100 bp
 * 81573 94653: contig of 13081 bp in length
 * 94654 94753: gap of 100 bp
 * 94754 106759: contig of 12006 bp in length
 * 106760 106859: gap of 100 bp
 * 106860 118633: contig of 11774 bp in length
 * 118634 118733: gap of 100 bp
 * 118734 135004: contig of 16271 bp in length
 * 135005 135104: gap of 100 bp
 * 135105 151780: contig of 16676 bp in length.
 Location/Qualifiers
 1..151780

FEATURES
 SOURCE

Db 113251 TCTGACAGCCAAAGATGAAACAAACAGAAAAAAGTAAAGACTTATTATGCT 113310
QY 1479 gacataattacggctgacaactccctggaagaagctatgcttccagcgtgctcc 1538
Db 113311 GACATATTTACGGCTACAAACTCTGGAGAGATGATGCTTCCAGCTGGCTCC 113370
QY 1539 ccgagatgttgctaccctccctccatctcaagaagaatacatcatccatgaggtta 1598
Db 113371 CCGGATGTTGGCTACCTCACCCTCCATCTCAAGAAATATCATCTCATTTGGGGA 113430
QY 1599 gaaaagagaaggtccgaggtgfgggaaggaatagaatcaatccgcccaactcc 1658
Db 113331 GAAAGAGAGAGGCTCCGAGGGGTGGAGGATGAAATACATCCGCCCAACTCC 113490
QY 1659 aaagagcagatccctcccccagccatagcatgttttaagtcacactccgaagaa 1718
Db 113491 AAAGAGAGATCCCTCCCGACCCATAGCATGTTTAAAGTCACCTTCCGAGAGAA 113550
QY 1719 gtaagaagtcagagacatgagcttgcaagccggaagagacatcaaaactca 1778
Db 113551 GTGAAAGTTCAGAGACACTGGCTTGCAGGCCCGAGGACACCATCACAACATCACA 113610
1779 gacacagacatccctcttgagacacgacctctgcccacacacacgagacatltctgc 1838
Db 113611 GACACAGACATCCCTTTGAGACACGCCCTTCTGCCACATCAGCAGACACATTTCTGC 113670
QY 1839 ctagaagaacgctcttactgctcttaacatgtagatatacttaacataaagaatat 1898
Db 113671 CTGAAAGAGCTTCTTACTGCTCTTACATGATGATGATGATGATGATGATGATGAT 113730
QY 1899 tactgggggaaaaaacatcacatgctgtaacatctgagaacacgagacataaagat 1958
Db 113731 TATTGGGGGAAAAACTACAAAGTGTACATATGCTGAGAAATGACAGACATATAT-CT 113789
QY 1959 gccacccaaaacactttggaatacatctccagaacaaccttactcttggtagtt 2018
Db 113790 GCCACCCAAAATCTTTTGAATAATCATTTCCAGAACACTCTTACTTCTGATGAT 113849
QY 2019 ttaactgttaaaaaaaagtttaaacagaagacacatgacatagaagcctgcagga 2078
Db 113850 TTAATGCTTAAAAAAGTTTAAACAGACACATGACATGATGATGATGATGATGATGAT 113909
QY 2079 ctgagcttttttggcaatcttcacagtggaactttccacaagaataaagatgag 2138
Db 113910 CTGAGCTTTTGTGGCAATTTCTCCAGCTGGACTTCTCCAGAAATGAAGTACTGG 113969
QY 2139 tttttaagaagctaaagttacatatttctcacttaagttacttaagcaaaagttt 2198
Db 113970 TTTTAAAGAGTTAACTATATTTATTTCTCACTTAACTTATTTATGCAAAAGCTTT 114029
2199 tctgttagagaaatgacaatgtaatatgctcttaagtaataacagctcgtctccagag 2258
Db 114030 TCTTTAGAGAAATGACAAATGTTAATTTGCTTATGAAATTAACAGTCTGTTCCAGAG 114089
QY 2259 tccaagacatgtaataaagacatgaatcatgacggaag 2301
Db 114090 TCCAGAGACATGTTTATTAAGACATGATCATGACCGAAG 114132
RESULT 7
AC073954 177744 bp DNA HTG 17-NOV-2000
LOCUS AC073954/c Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
DEFINITION SEQUENCE, 34 unordered pieces.
ACCESSION AC073954
VERSION AC073954.3 GI:10998993
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 177744)
AUTHORS Waterston, R.H.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

the sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 177744)
Waterston, R.H.
Direct Submission
Submitted (07-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Oct 25, 2000 this sequence version replaced gi:19186849.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----
Center project name: H.NH0147C10
----- Summary Statistics -----

Sequencing vector: Plasmid; 08
Chemistry: Dye-Primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 155679 bases at least Q40
Consensus quality: 163329 bases at least Q30
Consensus quality: 165780 bases at least Q20
Insert size: 18500; agarose-fp
Insert size: 183559; sum-of-contigs
Quality coverage: 3.29 in Q20 bases; agarose-fp
Quality coverage: 3.43 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1247: contig of 1247 bp in length
* 1248 1347: gap of unknown length
* 1348 2762: contig of 1415 bp in length
* 2763 2862: gap of unknown length
* 2863 4773: contig of 1911 bp in length
* 4774 4873: gap of unknown length
* 4874 6326: contig of 1453 bp in length
* 6327 6426: gap of unknown length
* 6427 8673: contig of 2447 bp in length
* 8674 8973: gap of unknown length
* 8974 10783: contig of 1810 bp in length
* 10784 10883: gap of unknown length
* 10884 12763: contig of 1880 bp in length
* 12764 12863: gap of unknown length
* 12864 15381: contig of 2518 bp in length
* 15382 15481: gap of unknown length
* 15482 17365: contig of 1884 bp in length
* 17366 17465: gap of unknown length
* 17466 19992: contig of 2527 bp in length
* 19993 20092: gap of unknown length
* 20093 23512: contig of 3420 bp in length
* 23513 23612: gap of unknown length
* 23613 25699: contig of 2087 bp in length
* 25700 28333: gap of unknown length
* 28334 28433: contig of 2534 bp in length
* 28434 32027: gap of unknown length
* 32028 32127: gap of unknown length
* 32128 36115: contig of 3988 bp in length
* 36116 36215: gap of unknown length
* 36216 39882: contig of 3667 bp in length
* 39883 39982: gap of unknown length
* 39983 44009: contig of 4027 bp in length
* 44010 44109: gap of unknown length
* 44110 47829: contig of 3720 bp in length

*	47830	47929:	gap of unknown length
*	47930	53768:	contig of 5833 bp in length
*	53769	53868:	gap of unknown length
*	53869	60215:	contig of 6347 bp in length
*	60216	60315:	gap of unknown length
*	60316	66477:	contig of 6162 bp in length
*	66478	66577:	gap of unknown length
*	66578	71476:	contig of 4899 bp in length
*	71477	71576:	gap of unknown length
*	71577	78883:	contig of 7307 bp in length
*	78884	78983:	gap of unknown length
*	78984	83952:	contig of 6965 bp in length
*	83953	86052:	gap of unknown length
*	86053	93510:	contig of 7458 bp in length
*	93511	93610:	gap of unknown length
*	93611	100718:	contig of 7108 bp in length
*	100719	100816:	gap of unknown length
*	100819	109795:	contig of 8977 bp in length
*	109796	109895:	gap of unknown length
*	109896	117930:	contig of 7893 bp in length
*	117931	117990:	gap of unknown length
*	117991	126044:	contig of 8134 bp in length
*	126045	126044:	gap of unknown length
*	126145	134442:	contig of 8298 bp in length
*	134443	134442:	gap of unknown length
*	134543	145663:	contig of 11121 bp in length
*	145664	145663:	gap of unknown length
*	145664	157189:	contig of 11426 bp in length
*	157190	157283:	gap of unknown length
*	157290	160822:	contig of 9533 bp in length
*	160823	166922:	gap of unknown length
*	166923	177944:	contig of 10822 bp in length

UNRES	Location/Qualifiers
source	1..177744
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="17"
	/clone="RP11-147C10"
misc_feature	1..1247
misc_feature	/note="assembly_name:Contig26"
misc_feature	1348..2762
misc_feature	/note="assembly_name:Contig31"
misc_feature	2863..4773
misc_feature	/note="assembly_name:Contig34"
misc_feature	4874..6326
misc_feature	/note="assembly_name:Contig35"
misc_feature	6427..8873
misc_feature	/note="assembly_name:Contig36"
misc_feature	8974..10763
misc_feature	/note="assembly_name:Contig37"
misc_feature	10884..12763
misc_feature	/note="assembly_name:Contig38"
misc_feature	12864..13581
misc_feature	/note="assembly_name:Contig39"
misc_feature	15482..17165
misc_feature	/note="assembly_name:Contig40"
misc_feature	17466..19992
misc_feature	/note="assembly_name:Contig41"
misc_feature	20093..23512
misc_feature	/note="assembly_name:Contig42"
misc_feature	23613..25699
misc_feature	/note="assembly_name:Contig43"
misc_feature	25800..28333
misc_feature	/note="assembly_name:Contig44"
misc_feature	28434..32027
misc_feature	/note="assembly_name:Contig45"
misc_feature	32128..36115
misc_feature	/note="assembly_name:Contig46"
misc_feature	36216..39682
misc_feature	/note="assembly_name:Contig47"
misc_feature	39983..44009
misc_feature	/note="assembly_name:Contig48"
misc_feature	44110..47829

```

misc_feature      /note="assembly_name:Contig49"
                  47930..33768
                  /note="assembly_name:Contig50"
misc_feature      53869..60215
                  /note="assembly_name:Contig51"
misc_feature      60316..66477
                  /note="assembly_name:Contig52"
misc_feature      66578..71476
                  /note="assembly_name:Contig53"
misc_feature      71577..78883
                  /note="assembly_name:Contig54"
misc_feature      78984..85952
                  /note="assembly_name:Contig55
clone_end:87
vector_side:left"
misc_feature      86053..93510
                  /note="assembly_name:Contig56"
misc_feature      93611..100718
                  /note="assembly_name:Contig57"
misc_feature      100819..109795
                  /note="assembly_name:Contig58"
misc_feature      109896..117790
                  /note="assembly_name:Contig59"
misc_feature      117891..126044
                  /note="assembly_name:Contig60"
misc_feature      126145..134442
                  /note="assembly_name:Contig61"
misc_feature      134543..145663
                  /note="assembly_name:Contig62"
misc_feature      145764..157189
                  /note="assembly_name:Contig63"
misc_feature      157290..166822
                  /note="assembly_name:Contig64"
misc_feature      166923..177744
                  /note="assembly_name:Contig65"

```

Query Match	46.3%	Score 1066.4	DB 75	Length 177744	
Best Local Similarity	97.0%	Pred. No. 1.7e-15			
Matches 1129	Conservative	0	Mismatches 31	Indels 4	Gaps 4
Qy 1142	aacgtatgagtgttctt- agltccygtctcggcac- taactgtcgtgtaacctgaa	1199			
Dy 93369	AGCTGGAGATGTGTGTGCTGATGTCCTGGCGTCCACTTACTTGTGGGAAATCCGGAA	93310			
Qy 1200	ctaacacattctctcttcggagcctaattcca- -cttggtaaatgaggttgaggtgag	1258			
Dy 93309	CTACACAGTGTGTGGCGGAGCCTCTCACTACTGTGTGTAATGATGGTGGAGGTGGG	93250			
Qy 1259	aatagatcttcgagagagactatgtgataatgata ttccaagagacccagtgccctttgaa	1318			
Dy 93249	AATGAGATCTCGAGAGACTATTGGCATTTGATGTGATCCAGAGACTCCAGTGCCTTTTGAT	93190			
Qy 1319	ggcagagttagagagagagagagagagagagagatgaatgaatgttcattgattcag	1378			
Dy 93189	GGCAGAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGATGCACTTGATTAATTCAG	93130			
Qy 1379	ggcagaatcacttccagaatt- cagaattgtgatagtctctcttcgcacgccaaagtga	1437			
Dy 93129	GCCAGGCTACTTCCACAAATTTNAGAGTTGTATCTCTCTTTCGACGCCAAAGATGAA	93070			
Qy 1438	aaacaaacagaaaaaaaaggtlaaagagctata ttatgtcgtgaataatlaacgctgaca	1497			
Dy 93069	AAACAAACAGAAAAAAGCTAAAGAGCTAATTATGGCTGACATATTATCCGCTGACA	93010			
Qy 1498	aactcttggaagaagctatgtctctctccagcctgtg cttcccgagatglttggtacttc	1557			
Dy 93009	AACCTCTGGAGAGAGCTATCTCTCTTCCAGCCTGGCTTCCCGGATGTTGGCTACCTC	92950			
Qy 1558	caaccctcactctlaaagaaataacatacattggtgtg tlaaagagagaggtccgag	1617			

D	92349	CACCCCTCCATCTCAAGAAGAAATACATCATTCATTGGGATGAGAAAGAGAGGGTCCGAG	92890
Q	1618	ggtgtgtggagggatagaataacatccgccccactttccaaagagagatccctccc	1677
D	92889	GGTGTGTGGAGGGATGAAATTCATCTCCGCCCAACTTCCAAAGAGAGCATCCCTCC	92830
Q	1678	ccgagcccatagccatgtttttaaagttaccttccgaagaaagtgaagtttcaagagcac	1737
D	92829	CCGACCATATGGCATTTTAAAGTTCACCTTCCGAAGAAAGTGAAGGTTCAGAGCAC	92770
Q	1738	tggcctctgagagcccgaaaggagacgacatcaaaactcaagaccagagactccctttg	1797
D	92769	TGGCTTTGCAGGCCCGAGGGAGAGGCATCACAAACTCACAGACGACATCCCTTTGG	92710
Q	1798	agacacgcgcctcttgcccaccactcaagagacatlttcgcctagaagaaacgcctcttac	1857
D	92709	AGACACGCGCTTTGGCCCACTCCAGACGACATTTCTGCTGTGAAACAGCTTTTAC	92650
Q	1858	tgcctcttaccatgtgatctggatctcttaccacttaagaaatattatttgggggaaactaca	1917
D	92649	TGCCTTTACATGTGATGGCATATCTTACCTTAAAGAAATATTATTGGGGAAACACTACA	92590
Q	1918	agttgctgtacatatgtctgagaaactctgcagagataatagctgtccaccccaaatctttc	1977
D	92589	AGTCTCTACATATGCTGTGAGAAACTCCAGACGATATATCTCCACCTCCAAATAATCTTTT	92530
Q	1978	gaatacatcttccagacaacctcttactcttctgtgttgattttaaattgttaaaaaaaa	2037
D	92529	GAAATATATTCCAGACAACTCTTACTTCTTCTGTGTAGTTTATTGTGTTTAAANAAAAA	92470
Q	2038	agttttaacagaaagcacatgtacatatgaagaactgcagagactgttcgttttttggcaa	2097
D	92469	AGTTTAAACAGAGACATGACATGTGAAGAACCTCGAGGACTGTGCTTTTGTGGCAA	92410
Q	2098	ttcttccacgtggagactctgtccaaagaatgaagaatgattgtttttaaaggttaagtta	2157
D	92409	TTCTTCCAGCTGGGACTTGTCCACAAGAAATGAATAGTGGTTTAAAGAGTTAACTTA	92350
Q	2158	catattatcttctcaacttaagattatattacaaagtttcttctgttagagaatgacat	2217
D	92349	CATATTTATTTTCTCACTTAACTGATTATTATTTGCAAAAAGTTTTCTTGTAAGAAATGACAAAT	92290
Q	2218	gttaatactgctttaagaaataacagctcgttcttccagagtcacagagacatgttata	2277
D	92289	GTTAAATATGCTTTATGAATTAACAGTGTGTCTTCAGAGTCCAGAGACTGTTTATA	92230
Q	2278	aagacaatgatcatgtaccgaaag	2301
D	92229	AAGCAATGAAATCATATGACGAAG	92206
LOCUS	AF326742	642 bp	mRNA
DEFINITION	Cercopithecus aethiops sclerostin mRNA, complete cds.		
ACCESSION	AF326742		
VERSION	AF326742.1	GI:13161028	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			

AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.M., Kovalevich,B.R., Prolli,S., Skonler,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alishch,R.S., Gillett,L., Colbert,T., Tacconelli,P., Galas,D., Hemerik,H., Beighton,P. and Mulligan,J.T.
TITLE	Direct Submision
JOURNAL	Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA
FEATURES	Location/Qualifiers
SOURCE	1..642 /organism="Cercopithecus aethiops" /db_xref="taxon:9534"
CDS	1..642 /codon_start=1 /product="sclerotin" /protein_id="AAK13457.1" /db_xref="gi:13161029"/> /translation="MQLPLALCVLCILVHAARFVEYGGQMAFKNDATETIIPLEGEP EPPLENNKTMMNRANGRPHPHPTKVSEYSCRELHFRTYDOPCSAKPVED LVSCGCGPARLLNPAIGRCKWMPSPGPDRCPIDRYRAQROLLCPGGAAPRAKYR LVAASKRLIRFNOSTLKDFGEARLRPKGRKPRFRARKAMDALENAY"
BASE COUNT	113 a 228 c 210 g 91 t
ORIGIN	
Query Match	26.6%; Score 613.2; DB 89; Length 642;
Best Local Similarity	97.2%; Pred. No. 5,1e-108;
Matches 624;	Conservative 0; Mismatches 18; Indels 0; Gaps 0;
OY	48 atgcagctcccaatgyccttgctgtcgtcgtcgtcgttaacacagccttcgtga 107
Db	1 ATGCAGCTCCCAATGCGCCCTGTGTCTTGTCTTCGCTGCTACACGACGCTTCCTGTGA 60
OY	108 gtggaggcgcaaggagggtgcagagcgltcaagaatgatgcacaaggaaacatcccagctc 167
Db	61 GTGAGAGGCCAGAGGGGTGGCAGGCGCTTCAAGAATGATGCCACAGAAATCATCCCCGAGCTC 120
OY	168 ggagagatraccccgagcctccaccgagccttggagaaacaagaacctgaaecggggcgag 227
Db	121 GGAGAGTACCCTCAGAGCTCTCACCGAGCTGGAGAACAAACAAGACCATGAACCGGGGAG 180
OY	228 aacgagaggcgactccccaccacccctttgagaccaagaagcgtlccagtagaacagtgc 287
Db	181 AATGAGAGGGGGCGCTCCCACACCCCTTTTGAGACAAAGACGATCGAGTACACTGC 240
OY	288 cggaagctgacttaaccgcgaactgagacgaatggccgttgcgcgacgacgcaagccggc 347
Db	241 CGAGAGCTGACTTACCCCGCTAAGTAGACCGATGGCCGTCGCGACGCGCAAGACCAATG 300
OY	348 accgagctgtgtgtcctcggcagatgacgagccggcgcgctcgtcgtcccaagcgcatcggc 407
Db	301 ACCGAGTTGGTGTGCTCCGGCCAGTGGCGGCCCGGACCCGCTGCTGCCAACGCAATCGGC 360
OY	408 cgagcgcaagtggtggagacactgtagggcggaatttcgctgtgcatcccgagccgctacggc 467
Db	361 CCGCGCAAGTGGTGGGCGCCGAGTAGGCCCATCTTCGCTGCATCCCGACCCCTACGCC 420
OY	468 gcgcagcgctgacagctgtcgtgtcccggttgtagagcgccgcgcgcgcgcgcgaagtgctgc 527
Db	421 GCGAGCGTGTGACGCTGCTGTGCCGATGTGCGCGCGCGCGCGCGCGCAAGGTGCGC 480
OY	528 ctgtgtggcctcggtggaagtgtcaagggccttaaccgccttcaaaaacagtctgaagctcaag 587
Db	481 CTGTGTGCTCTGTGCAAGTGAAGGCGCTCACCCGCTTCCACAACACAGTGGAGCTCAAG 540
OY	588 gaactcgagaccgaagccgctcgcgcgcgacagaagggccggaagccgcgcgcgcgcgcgg 647
Db	541 GACTTGGCTGCCAGGCGCCTCGCGCGCAGAAAGGCGGGAAGCCGCGCGCCCGCGCGG 600
OY	648 aggcgccaaagccaacagcgcgagcttggaagaagcctaactag 689
Db	601 GGGGCCAAAGCCATCAGCGCGAGCTGGAGAACCGCTTAAG 642

Db 475 CTGATGACCTGCTGCAAGTGAAGGCGCTCACCGCTTCCAAACCACTCGAGCTCAAG 534

Qy 588 gactctggagacgagccgctcggcgaggaagggcgagcgagcccgcccgccggcggc 647

Db 535 GACTTCGGGCGGAGACCGCGGCGGAGAGGGTTCGAGACCGCGCGCGCGCGG 594

Qy 648 agcgccaaagccacacgagcgagctgtagaagcctaagag 691

Db 595 GGAGCCAAAGCAACGAGCGGAGCTGGAGAACGCCCTACTAGAG 638

RESULT 12

AF326738 532 bp mRNA MAM 28-FEB-2001

LOCUS AF326738 Bos taurus sclerostin mRNA, partial cds.

DEFINITION AF326738

VERSION AF326738.1 GI:13161016

KEYWORDS

SOURCE COW.

ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

1 (bases 1 to 532)

Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene Product, a Novel Cystine Knot-Containing Protein

Am. J. Hum. Genet. 68 (3), 577-589 (2001)

11179006

JOURNAL 2 (bases 1 to 532)

PUBMED Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Direct Submission

Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

TITLE JOURNAL

FEATURES

source

1. 532

Location/Qualifiers

/organism="Bos taurus"

/db_xref="taxon:9913"

<L>.532

/codon_start=3

/product="sclerostin"

/protein_id="AAK13453.1"

/db_xref="GI:13161017"

/translation="MDAREIIPELGEYPEPLNKTNRNRAENGRRPHPEPTKDA SEYSCRELFHFRYYTDGCRSAKRYTELCSGGCGPARLLPNAIGKRRKRPSPFR CIPDRRAORVOLLCPGGAAPRAKRVLRVASCCKRLTRHNSSELKDFEPAARPOT GKKLRPRAGTKASNA"

CDS

BASE COUNT 95 a 199 c 176 g 62 t

ORIGIN

Query Match 20.3%; Score 467.8; DB: 7; Length 532;

Best Local Similarity 93.5%; Pred. No. 4.7e-80;

Matches 500; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

Qy 136 agaatgacgcaagaatcattcccgagctcgagagtagtaccgagctcaccgagc 195

Db 1 AGAATGATGCCACAGAAATCATCCCGAGCTGGCGAGTACCCGAGCTCTGCCAGAGC 60

Qy 196 tggagaaacaaagacatgaaacggcgagaaagggcgagcgctccaccaccct 255

Db 61 T--GAACAACAAGACCAATGAAACCGGGCGGAAGAGGGAGACTCCACCAACCCCT 117

Qy 256 ttgagcaaaagacgltccgagtagtagtgcgagagctgaacttaccgagcagc 315

Db 118 TTGAGACCAAGAGCGCTCGAGTACAGCTGCGGAGAGTGTGACTTACCCGCTACGGA 177

Qy 316 ccgatggccgtgtccgagcgagccaaagcagtcacagagctgtgtgtctccgacagtcg 375

Db 178 CCGATGGCCCTGTGCGCGAGCGCCAGCGGCTCACAGACTGAGTGTCTGGGCGAGTCG 237

Qy 376 gcccgagcgagctgtctgtccaaagcagtcagcgagcgaagtgtgtgtgtgtgtgtgt 435

Db 238 GCCCGCGCGCGCTCTGTGCGCGAGCGCCATCGCGCGCAAGTGTGTGTGTGTGTGTGT 297

Qy 436 ccgacttcgctgtatcccgagccgagcagcagcgagcagcgagcagcagcagcagcagc 495

Db 298 CCGACTTCGCTGTGATCCCGACCGCTACCGCGCGAGCGGGGTGAGCTGTGTGTGTGTGT 357

Qy 496 gt 555

Db 358 GCG 417

Qy 556 tcaccgcttcacaaacagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 615

Db 418 TCACTCGCTTCCACACCACTCCGAGCTCAAGACTTCGGGCGCGCGCGCGCGCGCGCG 477

Qy 616 aagaagcgaggaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 670

Db 478 AACCGCGCGGAGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 532

RESULT 13

AF326737 17423 bp DNA ROD 28-FEB-2001

LOCUS AF326737

DEFINITION AF326737

ACCESSION AF326737

VERSION AF326737.1 GI:13161013

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 17423)

Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene Product, a Novel Cystine Knot-Containing Protein

Am. J. Hum. Genet. 68 (3), 577-589 (2001)

11179006

JOURNAL 2 (bases 1 to 17423)

PUBMED Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Direct Submission

Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

TITLE JOURNAL

FEATURES

source

1. 17423

/organism="Mus musculus"

/strain="129Sv/J"

/db_xref="taxon:10090"

/chromosome="11"

/join(8036..8249,10742..11163)

/product="sclerostin"

/join(8036..8249,10742..11163)

/codon_start=1

/product="sclerostin"

/protein_id="AAK13452.1"

/db_xref="GI:13161014"

/translation="MQPSLAPLILULYHAACFAVEGGQMAFNDATVYIPGEGP EPPENNQTNNRAENGRRPHHPYPAKDVSEYSCRELFHFRYYTDGCRSAKRYTELV CSQCGPARLLPNAIGRVKWRPNPDRCTIPDRRAORVOLLCPGGAAPRVLY ASQCKRLTRFHNSSELKDFETARPQKGRKPRGARGAANAAELENAY"

CDS

BASE COUNT 4456 a 4306 c 4569 g 4092 t

ORIGIN

Consensus quality: 181484 bases at least Q40
Consensus quality: 190159 bases at least Q30
Consensus quality: 193745 bases at least Q20
Insert size: 207000; agarose-fp
Insert size: 197427; sum-of-contigs
Quality coverage: 3.8 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1091: contig of 1091 bp in length
1092 1191: gap of 100 bp
1192 3065: contig of 1874 bp in length
3066 3165: gap of 100 bp
3166 4305: contig of 1140 bp in length
4306 4405: gap of 100 bp
4406 5983: contig of 1578 bp in length
5984 6083: gap of 100 bp
6084 8322: contig of 2239 bp in length
8323 8422: gap of 100 bp
8423 10098: contig of 1676 bp in length
10099 10198: gap of 100 bp
10199 12592: contig of 2394 bp in length
12593 12692: gap of 100 bp
12693 14684: contig of 1992 bp in length
14685 14784: gap of 100 bp
14785 17163: contig of 2379 bp in length
17164 17263: gap of 100 bp
17264 19900: contig of 2637 bp in length
19901 20000: gap of 100 bp
20001 23054: contig of 3054 bp in length
23055 23154: gap of 100 bp
23155 26011: contig of 2857 bp in length
26012 26111: gap of 100 bp
26112 29412: contig of 3301 bp in length
29413 29512: gap of 100 bp
29513 32991: contig of 3479 bp in length
32992 33091: gap of 100 bp
33092 36913: contig of 3822 bp in length
36914 37013: gap of 100 bp
37014 40880: contig of 3867 bp in length
40881 40980: gap of 100 bp
40981 45015: contig of 4035 bp in length
45016 45115: gap of 100 bp
45116 50321: contig of 5206 bp in length
50322 50421: gap of 100 bp
50422 54284: contig of 3863 bp in length
54285 54384: gap of 100 bp
54385 59430: contig of 5466 bp in length
59431 59530: gap of 100 bp
59531 64373: contig of 4843 bp in length
64374 64473: gap of 100 bp
64474 70503: contig of 6030 bp in length
70504 70603: gap of 100 bp
70604 77829: contig of 7226 bp in length
77830 77929: gap of 100 bp
77930 85976: contig of 8047 bp in length
85977 86076: gap of 100 bp
86077 91694: contig of 5618 bp in length
91695 91794: gap of 100 bp
91795 99208: contig of 7414 bp in length
99209 99308: gap of 100 bp
99309 107499: contig of 8191 bp in length
107500 107599: gap of 100 bp
107600 115731: contig of 8132 bp in length
115732 115831: gap of 100 bp
115832 123297: contig of 7466 bp in length

123298 123397: gap of 100 bp
123398 133579: contig of 10582 bp in length
133580 134079: gap of 100 bp
134080 144631: contig of 10552 bp in length
144632 144731: gap of 100 bp
144732 159118: contig of 14387 bp in length
159119 159218: gap of 100 bp
159219 172437: contig of 1319 bp in length
172438 172537: gap of 100 bp
172538 200727: contig of 28190 bp in length.

Location/Qualifiers
1. .200727
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-346P7"
/clone.lib="RPC1-23 female Mouse BAC"
1. .1091
/note="assembly-fragment"
1192. 3065
/note="assembly-fragment"
3166. 4305
/note="assembly-fragment"
4406. 5983
/note="assembly-fragment"
6084. 8322
/note="assembly-fragment"
8423. 10098
/note="assembly-fragment"
10199. 12592
/note="assembly-fragment"
12693. 14684
/note="assembly-fragment"
14785. 17163
/note="assembly-fragment"
17264. 19900
/note="assembly-fragment"
20001. 23054
/note="assembly-fragment"
23155. 26011
/note="assembly-fragment"
26112. 29412
vector_end:SP6
/note="assembly-fragment"
29513. 32991
/note="assembly-fragment"
33092. 36913
/note="assembly-fragment"
37014. 40880
/note="assembly-fragment"
40981. 45015
/note="assembly-fragment"
45116. 50321
/note="assembly-fragment"
50422. 54284
/note="assembly-fragment"
54385. 59430
/note="assembly-fragment"
59531. 64373
/note="assembly-fragment"
64474. 70503
/note="assembly-fragment"
70604. 77829
/note="assembly-fragment"
77930. 85976
/note="assembly-fragment"
86077. 91694
/note="assembly-fragment"
91795. 99208
/note="assembly-fragment"
99309. 107499
misc_feature

```

/note="assembly_fragment"
misc_feature
107600..115731
/note="assembly_fragment"
misc_feature
115832..123297
/note="assembly_fragment"
clone_end:r7

```

misc_feature

107600. .115731

3
2
1
0
9
8
7
6
5
4
3
2
1
0

```
/note="assembly_fragment"
```

misc_feature

```
clone_end:T7
```

clone_end:T7

Query Match

16.2%; Score 372.2; DB 62; Length 200727;

Best Local Similarity 85.9%; Pred. No. 5.2e-62;
Matches 413; Conservative 0; Mismatches 59; Total 472

Matches 413; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

233 agggcggtctcccaaccaaccttgagaccaaagacgtgtccgaglacagctgccgca 292

Db 150286 AGGCTGCAGCCCCCTCAGCATCCCTCTCTCCGAGATGTGTCCGAGTACAGCTGCCCGGA 150345

293 gctgcactcaaccgctacgtgacgatatggccgttgccgacgccaagccggtcaccca 352

150346 GCTGCACTACACCCGCTTCCTGACAGACGGCCCATGCGCGAGCGCCAAGCCGGTCACCGA 150405

353 gctggtgtgctccgcgcagtcgcgcgcgcgcctgtctgcccaacgccatcgccgcg 412

Db 150406 GTGGTGCTCCGGCCAGTGGGGCCCCGGGGCTGCTGCCCAACGCCATCGGGCGGT 150465

413 caagtgtgacgacctagtggccgacttcgctgcattcccgaccgctacgcgcga 472

Db 150466 GAAGTGGTGGCCCGAACGACCGGATTTCCGCTGCATCCCGGATCGCTACCGCGCGCA 150525

QY 473 gcgcgtgcaagctgtgtgtcccggtgtgtagggcgcgcgcgcgaagtgcgcctgt 532

Db 150526 GCGGGTGCAGCTGCTGTGCCCGGGGGCGGGCGCGGCTCGCGCAAGTGCGTCTGGT 150585

533 ggcctcgtgcaagtgcgaagcgcctcacccgcttcacacaaccagtgcgagctcaagactt 592

Db 150586 GGCCTCGTGCAAGTCAAGCGCCTCACCCGCTTCACACAACCAAGTCCGAGCTCAAGACTT 150645

593 cgggaccgagccgctcggccgcagaaaggccggaagcccgccgagcgc 652

Db 150646 CGGGCCGAGACCGCGCGCCGAGAGGGTGCACAAGCCGCGGCCCGCGGCCCGGGAGC 150705

653 caagccaaccagccgagctggagaacgctactagagccgcgcgcctccac 712

Db 150706 CAAGCCAACCGAGCGGAGCTGGAGAACGCCCTACTAGAGCGAGCCCGCGCCTATGCAGCC 150765

QY 713 c 713

150766 C 150766

Search completed: November 8, 2001, 23:49:26
Job time: 16136 sec

Job time: 16136 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:49:26 ; Search time 9526.57 Seconds
(without alignments)
3736.003 Million cell updates/sec

Title: US-09-668-021-5
Perfect score: 2301
Sequence: 1 agagcctgctactcgaag.....caatgatcatgatccgaaag 2301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

rchcd: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBml: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pal1: *
10: gb_pal2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_v1: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v1: *
59: gb_v12: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_pr1: *
95: gb_pr2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2297.8	99.9	2323	89	AF326739 Homo sapi
2	2277.4	99.0	2329	9	AX056687 Sequence
3	2270	98.7	2296	89	AF331844 Homo sapi
4	2047.8	89.0	21501	85	AF326736 Homo sapi
5	2047.8	89.0	94752	85	AC003098 Homo sapi
6	2035.8	88.5	151780	72	AC055813 Homo sapi
7	1066.4	46.3	177744	75	AC073954 Homo sapi
8	610	26.5	642	89	AF326742 Cercopit

9	53.3	23.2	674	94	AF336741	AF336741 Rattus norvegicus
10	51.3	22.3	177744	75	AC073954	AC073954 Mus musculus
11	50.2	21.8	638	94	AF326740	AF326740 Mus musculus
12	466.2	20.3	532	7	AF326738	AF326738 Bos taurus
13	372.2	16.2	17423	94	AF336737	AF336737 Mus musculus
14	372.2	16.2	110000	72	AC068788-2	Continuation (3 of AC012266 Mus musculus
15	372.2	16.2	200727	62	AC012266	AC012266 Mus musculus
16	372.2	16.2	211533	73	AC068807	AC068807 Mus musculus
17	158	6.9	51575	68	AC023810	AC023810 Mus musculus
18	74.4	3.2	204242	77	AC087891	AC087891 Mus musculus
19	69.6	3.0	900	9	AF017832	AF017832 Sequence
20	64.4	2.8	150951	75	AC079155	AC079155 Homo sapiens
21	62.8	2.7	1500	93	HS0800108	AL050024 Homo sapiens
22	61.6	2.7	229380	76	AC079636	AC079636 Mus musculus
23	61.2	2.7	162005	80	AL356002	AL356002 Homo sapiens
24	61	2.7	567	53	CNS07541	AL429856 Clone BAO
25	60.6	2.6	38911	86	AC005559	AC005559 Homo sapiens
26	60.2	2.5	63081	67	AC022663	AC022663 Homo sapiens
27	58.4	2.5	165448	61	AC010827	AC010827 Homo sapiens
28	58	2.5	4523	59	MM010GENS	X95710 pseudoviral
29	57.8	2.5	412	10	AX071502	AX071502 Sequence
30	57.8	2.5	11214	1	AE005002	AE005002 Halobacterium
31	56.5	2.5	159968	77	AC084805	AC084805 Mus musculus
32	56.2	2.4	187738	71	AC027682	AC027682 Homo sapiens
33	56	2.4	33820	3	SE020	AL136058 Streptomyces
34	56	2.4	252859	68	AC024907	AC024907 Homo sapiens
35	54.8	2.4	217412	68	AC024400	AC024400 Homo sapiens
36	54.2	2.4	76734	67	AC023218	AC023218 Homo sapiens
37	54	2.3	567	53	CNS07541	AL429856 Clone BAO
38	53.8	2.3	17856	3	SKR20	AL450450 Streptomyces
39	53.8	2.3	52884	73	AC068018	AC068018 Homo sapiens
40	53.8	2.3	222426	76	AC084055	AC084055 Mus musculus
41	53.8	2.3	241271	78	AC090879	AC090879 Mus musculus
42	53.6	2.3	70097	68	AC023530	AC023530 Homo sapiens
43	53.4	2.3	51914	74	AC068022	AC068022 Homo sapiens
44	53.2	2.3	6876	89	AF335591	AF335591 Homo sapiens
45	53.2	2.3	45396	3	SC035	AL078610 Streptomyces

ALIGNMENTS

[illegible]

ORGANISM	TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS	FEATURES
Homo sapiens						source
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
1 (bases 1 to 2123)						
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeper,B.W.,						
Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,						
Fu,Y.H., Alish,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.						
Hamersma,H., Beighton,P. and Mulligan,J.T.						
Bone Dysplasia Sclerosteosis Results from loss of the SOST Gene						
Product, a Novel Cystine Knot-Containing Protein						
Am. J. Hum. Genet. 68 (3), 577-589 (2001)						
11179006						
2 (bases 1 to 2123)						
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeper,B.W.,						
Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,						
Fu,Y.H., Alish,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.						
Hamersma,H., Beighton,P. and Mulligan,J.T.						
Direct Submission						
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 16330						
220th St. SE, Bothell, WA 98021, USA						
location/Qualifiers						
1..2123						
/organism="Homo sapiens"						

Query Match	Best local Similarity	Score	DB	Length
Matches 2299; Conservative	99.9%; Pred. No. 0;	2297.8;	89;	2323;
	0; Mismatches	2;	Indels	0; Gaps
QY	1	aggagcctggtcctacggaagtggcgctgcccctctctgctggtgtaccatgacgtccac	60	
Db	1	AGAGCCTGTGTACTAGGAAGGTGGGTGCTCTCTGTGCTGTACCATGCACTCCAC	60	
QY	61	tggccctgtctcaatctgctctgctgctgtaacaacagcttcgctgtagtggagggccag	120	
Db	61	TGGCCTGTGTCTCGTGTGCTGCTGTGTACACAGGCTTCCGTATGTGAGGGCCAG	120	
QY	121	ggtgacagcgcttcaagaatgatgccagaaatcatccgagctccgagagtagaccg	180	
Db	121	GGTGGCAGCGCTTCAAGATGTGCGACAGAAATCATCCCGAGCTCGGAGTACCCG	180	
QY	181	agcttcacccgagctgtagaacaacagacatgtaaccgagcgaggaacgaagagcgac	240	
Db	181	AGCTTCACCCGAGCTGTGAGAACACAGACATGAAACGGCGGAGAAACGAGGGCG	240	
QY	241	ctccccaacacccctttagagccaagaacgtgtccagtaacgctgctgcgagctgact	300	
Db	241	CTCCCAACACCCCTTTGAGAACAAAGACGTGTCGAGTACACTGCGCGAGCTCCACT	300	
QY	301	tcaaccgctacgttgacaaatggcgctgcgcaagcgcaagcggtcaaccgactgtgt	360	
Db	301	TCAACCGCTACGTGACCAATGGCGCTGCGCAAGCGCAAGCGGTCAACCGACTGTGT	360	
QY	361	gctccgagcaagtgcgagcccgcgctgctgtcccaacgcatatggcgcgagcaagtgt	420	
Db	361	GCTCCGAGCAAGTGCAGGCCCGCGGCGCTGTGCGCCAAACGCATGCGCGGCAATGTGT	420	
QY	421	ggcgacctagtggtggccgacttcgctgtcatcccgagcgcttcagcgagcggtgc	480	
Db	421	GGCGACCTAGTGGTGGGCCCGCACTTCCTCATCTCCGACCGCTACCGGCGCAACGCTGC	480	
QY	481	agctgtctgtctccggtgtgtagagcgcgcgcgcaaggtgctgtagtgcctgt	540	
Db	481	AGCTGTCTGTCTCCGGTGTGTAGAGCGCGCGCGCAAGGTGCTGTAGTGCCTGTGT	540	
QY	541	gcaagtgcgaagcgctcacccgcttccaaacagctcgagctcaagactctgggacg	600	
Db	541	GCAAGTGCAGAGCGCTCACCCCTTCCAAACAGTGGAGCTCAAGGACTTGGGACCG	600	
QY	601	agggcgtcgcgcgaggaagggcggaagcgcgagcccgcccgccggaagcgcaaacgca	660	
Db	601	AGGCGCTCGCGCGAGGAAGGGCGGAAGCGCGGCCCGCGCCGAGGCGCAAAAGCA	660	
QY	661	accaagcgagctgtagaagcctctctagagcccgcgcgccctcccaaccgcgggc	720	
Db	661	ACCAAGCGAGCTGTAGAGCCCTCTAGAGCCCGCGCGCCCTCCACACGCGGGCG	720	
QY	721	ggccggccctgaaacccgcccccaattctgttcctctgagcggtgtgtgtgtat	780	
Db	721	GGCCGGCCCTGAACCCCGCCCCAATTCTGTCTCTGTGGCGTGTGTGTGTAT	780	

RESULT 3
AF331844 2296 bp mRNA PRI 06-MAR-2001
LOCUS Homo sapiens SOST (SOST) mRNA, complete cds.
DEFINITION AF331844
ACCESSION AF331844.1 GI:13236417
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Balemans, W., Ebeling, M., Patel, N., van Hul, E., Olson, P.,
Dioszegi, M., Lacza, C., Wuyts, W., van den Ende, J., Williams, P.,
Paez-Alves, A.F., Hill, S., Bueno, M., Ramos, F.J., Tacconi, P.,
Dijkers, F.G., Stratakis, C., Lindpaintner, K., Vickery, B.,
Foerzler, D. and Van Hul, W.,
Increased bone density in sclerosteosis is due to the deficiency of
a novel secreted protein (SOST)
Hum. Mol. Genet. 10 (5), 537-543 (2001)
1181578
JOURNAL
PUBMED
2 (bases 1 to 2296)
REFERENCE
AUTHORS
Balemans, W., Ebeling, M., Patel, N., Vickery, B., Foerzler, D. and Van
Hul, W.,
Direct Submission
Submitted (22-DEC-2000) Medical Genetics, University of Antwerp,
Universiteitsplein 1, Antwerp B2610, Belgium
LOCATION/Qualifiers
1. .2296
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q12-q21"
1. .2296
/gene="SOST"
38. .679
/gene="SOST"
/note="mutations may cause sclerosteosis"
/codon_start=1
/product="SOST"
/protein_id="AAK16158.1"
/db_xref="GI:13236418"
/translation="MOLPLALCLVLCILVHTAFRVVEGGQWAFKNDATTEILPELEGEP
EPPPELENNKTKNNRAENGSRPHEFERKDVSEYSCREIHFRTVTDPCRSARPYE
LYVSGGCGPARLLPNAIGRKMMRSPGPFNCIPDPRYAQRVQLCPGEGEAPRAKRV
LVASCKRRLTERHNSSELDKGTETEARPQGRKPRPARSAKANQALENAV"
GENE COUNT 576 a 631 c 607 g 482 t
GCIN

Query Match 98.7%; Score 2270; DB 89; Length 2296;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2284; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 11 ctatgtgaagtggtggtgcctctctgtgtgtaccatgcaactccactgtgctgtg 70
DB 1 CTACTGGAAGGTGGCGTCCCTCTCTGCTGTACATGCACTCCACTGACCTGTG 60
QY 71 tctaatctgctgtgtgtacacacagctcgcgtgtgtgtgtgtgtgtgtgtgtgt 130
DB 61 TCTGCTGCTGCTGCTGTGTACACAGGCTTCCGTGTAGTGTGAGGCGCAGGGGTGCGCAGGC 120
QY 131 gttcaagaatgtatgtgcacggaatcatcgcgagctcgcgagaggtaccctcgagctcacc 190
DB 121 GTTCAAGAATGTATGTCACGGAATATCATCCCGAGCTCGGAGATACCCCGACGCTTCACAC 180
QY 191 ggaagctggaagacacacagacatgtacacgagcggaagaaacgagagggcgctcccccacca 250
DB 181 GGAGCTGTGAGAACACAAAGCATGATGACCGGGCGGAGAAAGAGGGCGGCTCTCCACACA 240
QY 251 ccccttggagacgaagagctgttcgaggtacagctcgcgagctgtcaactcaaccgctta 310
DB 241 CCCCTTGGAGACCAAAAGAGCTGTCCGAGTACAGCTGCCGCGAGAGCTGCACCTTCACCGCGTA 300
QY 311 cgtgaccgatgggcccgtgcccgcagcgccgagccggtaccagagctgtgtctccgcgcca 370
DB 301 CGTATACGATGGGCCGTGGCCGAGCCGCAAGCCGGTACACGAGCTGTGTCTCTCCGGGCA 360
QY 371 gtcgagccgagcgccgtctgtctcccaacgcattcgcgcgagcaagtggtgtgcaactag 430
DB 361 GTGCGGCCCGCGCGCTGCTGTGCGCAACGCAATCGGCGCCGCGCAAGTGTGTGCGCACTAG 420
QY 431 tggagccgacttcgcgtgtgcatcccgaccgctaccgagcgagcgagcggtgtgcaactctgt 490
DB 421 TGAGCCCGCACTTCGCTGTGATTCGCCACCGCTACCGCGCGACGCGCTGACCTCTGTG 480
QY 491 tcccgtgtgtgagcgccgc 550
DB 481 TCCCGGTGTGAGCG 540
QY 551 ggcgctaccgctgtccacacacgaatcgaagctcgaagctcgaagctcgaagctcgaagctcga 610
DB 541 GCGCTCACCCGCTTCCACACACGATCGAGCTCAAGAGACTTGGAGACCGAGGCGCGCTCG 600
QY 611 ggcgcaagagcgccgagagcgccgc 670
DB 601 GCCCGAGAGAGCGCGGAGAGCG 660
QY 671 gctgtgagaagcgctactagaagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 730
DB 661 GCTGTGAAGACGCTACTATGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 731 tgaacccgcgcgcgcacattctgtctctgtgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 790
DB 721 TGAACCG 780
QY 791 aaatgtcctgcaaccgagcgagcgaggggtgtgagacttcgagccctgtgaggaatcccgagcg 850
DB 781 AAATGCTGTGCAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 851 ccgagcaagggccccctcagcccgcccgcaagctgaaagggtgcccaagggagggaggaatgt 910
DB 841 CCGGCAAGGCGCCCGCTTACG 900
QY 911 agagtcacagacactgtgacacgacgagcccgccctcgtggggcgcgactctgtgtgtgtc 970
DB 901 AGAGTCAACAGCACTGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 960
QY 971 cacttcagagagagcagaagaatgtgaagcatttcaacgcccctgaggggtgtttaaaggaagcgt 1030
DB 961 CACTTCAGAGAGGCGCAAAATGGAAGCATTTTCAACCGCCCTGGGGTTTAAAGGAGCGGT 1020
QY 1031 gtcggaagtggaaggtcccaagagctgtgttaagaagttgataaagattccccccttgacc 1090
DB 1021 GTGCGAGTGGGCAAAAGTCCACGAGCATGCTGTAAAGATTTGGATTAAGATTTCCCTTGCAC 1080
QY 1091 tctgtcccatcagaagaagccttgagggcggtgcccagagacacaaagccttgagggaacgtgtaga 1150
DB 1081 TCGCTGCCCATCAGAAAGCCTGAGAGCGTGCACAGACACAAAGACTG6GGGCAACTGTAGA 1140
QY 1151 tgtgtgttctagtctgtgtccttgcacaaactgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1210
DB 1141 TGTGTTTCTAGTCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1211 tccctggagacctcaatttccacttgtttaaagtgaaggtgtggaggtgtggaatgtgatactcg 1270
DB 1201 TCCCTGGAGACTCAATTTCCACTTTTAAATGAGAGGTGAGGTGAGATAGAGATCTCG 1260
QY 1271 agagactatgtgacatalgtattccaagagctccagtgcccttltgaaatggcgagaggtgag 1330
DB 1261 AGGAGACTATGTGGCATATGATATCCAAAGACTCCAGTCCCTTTGAATGGCGAGAGGTGAG 1320
QY 1331 agagagagagagagagagagagagagatgtatgtcaattgtatgtatgtatgtatgtatgtatgt 1390
DB 1321 AGAGAGAGAGAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380

[illegible]

ORGANISM	Homo sapiens Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.									
REFERENCE	1 (bases 1 to 21501)									
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alish,R.S., Gillett,L., Colbert,T., Tacconl,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T. Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene Product: a Novel Cysteine Knoc-Containing Protein									
TITLE	Am. J. Hum. Genet. 68 (3), 577-589 (2001)									
JOURNAL	11/19/06									
PUBMED	2 (bases 1 to 21501)									
REFERENCE	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alish,R.S., Gillett,L., Colbert,T., Tacconl,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.									
AUTHORS	Direct Submission Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA									
TITLE	Location/Qualifiers									
JOURNAL	1.. 21501									
FEATURES	source									
source	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="1/q21" join(<11961..12180,14939..>15360) /product="sclerostin" join(11961..12180,14939..15360) /codon_start=1 /product="sclerostin" /protein_id="AAK13451.1" /db_xref="GI:13161011" /translation="MDPLALCVLLVHTAFRVEVGQGAQFNKDATEIIPELGEP EPPLENNKTKMRAENGRRPHPEPTKXVSYSQRELFHFRVYTDPCRSAPVE IVASCCGPARLLPNAIGRKMWRSPGPRFCIPDPRARQVOLLCPGSEADPRARV IVASCKRLTRFHNQSELDKDFEAGRAPQGRKPRPARSKANQAELENAV"									
BASE COUNT	5546 a 5259 c 5419 g 5277 t									
ORIGIN										
Query Match	89.0%; Score 2047.8; DB 89; Length 21501;									
Best Local Similarity	98.9%; Pred. No. 0;									
Matches 2061; Conservative	0; Mismatches 22; Indels 0; Gaps 0;									
OY	219	cgggcggagagaacggagagcgccctcccaacacaccccttgagaccaaagcgtgtccgag	278							
DB	14890	CAGGGGGGGCAAGCGGGGGGCGACCCCTACCGCGGCTCTCTCAAGACGtGTCCGAG	14949							
OY	279	tacagctgcgcggagcttgtaactcaaccgcgtacgtgacgaatggccgttcgcgacgcgc	338							
DB	14950	TACAGCTGCGCGCAGCTGCACCTTCACCGGTACGTACGACCATGGGCGCGTCCGACGCGC	15009							
OY	339	aagccggtcaecgagcttggltgtctcggccaggtgagccggcgcgcctctgtcccaac	398							
DB	15010	AAGCGGCTCACCGAGTGGTGTGCTCCGGCCAGTGCAGCCCGCGCGCTGTGCGCAAC	15069							
OY	399	ggcatcgccgcggagcaaatggtgtggagcaactagtggcccgacatccgcgtgcatcccgac	458							
DB	15070	GCCATCGCGCGCGCGCAGTGTGTGGACACTTACTGGGCCCGACTTCGCTGCATCCCGAC	15129							
OY	459	cgcctaacgcgcgagcgcgtgtgacgttctgttcccggttggtggaagcgcgcgcgcgcgc	518							
DB	15130	CGCTACCGCGCGCAGCGCGTGCAGCTGCTGTCCCGGTGTGAGGCGCGCGCGCGCGC	15189							
OY	519	aaggtgtgcctcgtgtgcttcctgtgaacgtgacaagcgccctaacccgcttcacaaacagttcg	578							
DB	15190	AAGGTGTGCGCTGGTGTGCTCTGTGCACAGTGCACAGCGCCCTCACCGGCTTCACACCACTGG	15249							
OY	579	gagctcaaggaactctggagaccagagccgctcgcgcgcgacagaagggccggaagccgagccc	638							
DB	15250	GAGCTCAAGGACTCTTCGGGACCGAGGCGCGTGGCGCGGACGAAAGGGCCGGAAGCGCGGCCCC	15309							

OY 639 cgcgccgagcgccaaagccaaagccagcgagctggaagaagccctactagaagccgccc 698
 DB 15310 GCGCCCGGAGCGCGCAAGGCAACGAGCGCTGGAGAAACGCTACTACAGCCGCGCCC 15369
 OY 699 ggcgcctcccccacgcggcgagcgcccgccctgaaaccccgcccaattctgtctcc 758
 DB 15370 GCGCCCTCCCGACCGCGGGCGCCCGCTGAAACCGCGCCCGCAATTTCGTCTCC 15429
 OY 759 tgcgcgtgattgattgattatataatcattgtaaatgctgcgaaccccgagggcgagct 818
 DB 15430 TGCCGCTGGTGTGATGTATATTCATTGTAAATGCTGCAACCCGAGGCGAGGCGCT 15489
 OY 819 gaaagcttcagagccctcgaagaatcccgagcgagcgagcgagcgagcgagcgagcgagc 878
 DB 15490 GAGACCTTCAGAGCCCTGAGAGAAATCCCGGCGCGGCAAGCGCCCGCTACAGCCGCGCAGC 15549
 OY 879 tgaagggtcccaagggaggaagggaatgagagtcacagacactgagccagcgagcc 938
 DB 15550 TGAAGGGGTCCACGCGGCGAGGGGAATTCAGAGTCAACAGACCTGACGCGAGCGC 15609
 OY 939 ccgcctctggggcgccctactctgtgctccactcagaagggagcgagaaatgagagca 998
 DB 15610 CCGCCTCGGGGCGCGCTACTCTTCTGGTCCACTTCAGAGAGGCGAGAAATGGAAGCA 15669
 OY 999 ttctcacccgctgggggttttaaggagcggtggtggaggtgggaagttccagggactggt 1058
 DB 15670 TTTTCACCGCCCTGGGGTTTAAAGGAGCGGgtGGGAGTGGGAAGTCCAGGGAAGTGGT 15729
 OY 1059 taagaagttggaataagattccccccttgacactcgctgccatcagaagcgctgaagcgt 1118
 DB 15730 TAAAGAAAGTTGATAAAGATATCCCTTGCACCTCGCTCGCCATCGAAGGCGCTGAGCGCT 15789
 OY 1119 gcccaagcgcaaaagactggggggaactgtgagatgtgtgttctcaagtcctggctgcgacact 1178
 DB 15790 GCCGAGGACACAGACACTGGGGGCAACTGTAGATGTGCTTTCAGTCCGCTCGCTCGCAGCT 15849
 OY 1179 aactgtgtgtaaccttgactgaactacaactctctctcgagagctcaattccacttgt 1238
 DB 15850 AACTTGTGTGTAACTTGAACATACAAATTCCTTCGGGACCTGCAATTCCTTTCACCTTGT 15909
 OY 1239 aaaaatgaggtgaggtgaggaataagatctcgaagagagactatgtgcatatgattccaaag 1298
 DB 15910 AAAATGAGGGGTGGAGTGGGAATAGATCTCGAGGAGACTTATGTCATATGATTCGAAG 15969
 OY 1299 aactcagtgacctttgaatgagcgagaggtgagagagagagagagagagagagagagagaa 1358
 DB 15970 ACTCAGGCGCTTTTGAATGGGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 16029
 OY 1359 tgcagttgacttgattcaagtcgaaggtacttccagaattcgaagttgatagtctct 1418
 DB 16030 TGCAAGTTGCATTTGATTCAGTGCAGAGGTGCAAGTTCAGAAATTCAGAGTGTGATGCTCTCT 16089
 OY 1419 tctacagacccaagaatgaagaacaaacagaagaagaagaagaagaagaagaagaagaagaaga 1478
 DB 16090 TCTGACAGCCAAAGATCAAAAAACAAACAGAAAAAAGTAAAGAGTCTTTTATGTGCT 16149
 OY 1479 gaacataattacgctgacaaactcctcgagaagaagctagctgcttccagcctggtctcc 1538
 DB 16150 GACAAATTTACGGCTGACAAATCTCGGAGAGAGCTATGCTCTCTCCAGCTTGCTGCTCC 16209
 OY 1539 ccgagatgtttgactaccacacccctccatctcaagaagaataacatacattcaggggta 1598
 DB 16210 CCGGATGTTTGGTACTCAACCCCTTCATTCAAAGAAATTAACATTCATTCATTCGGGTA 16269
 OY 1599 gaaagagagagaggtcgaagaggtgtgtggagagatagaatccatccgcccccaactccc 1658
 DB 16270 GAAAGGAGAGAGGTCCAGAGGTGTGTGGAGAGATGAATTCATTCATTCATTCATTC 16329
 OY 1659 aagaagcagcatccctcccgagccacatagccaatgttttaagttacattccagaagaa 1718
 DB 16330 AAAGGCGAGCATCCCTCCCGAGCCATGAGCCATGTTTAAAGTCCACCTTCGGAAGAGAA 16389
 OY 1719 gtgaaggttcaaggaacatgctcttcgagcgccgagggagagagcatcacaactaca 1778

DB 16390 GTGAAAGTTCAAGAGACTGCGCTTCGAGGCCGAGGAGACAGCATCAAAACTGACA 16449
 OY 1779 gaccagacatcccttttgaagaccgctctgcccacactcaaggaagaacttctgc 1838
 DB 16450 GACCGACATCCCTTTTGAAGACCGCGCTGTGCGCCACGATCAGGAGACATTTCTCC 16509
 OY 1839 ctagnaacacgcttcttactgctcttactatgltgagtgatatacttactaactaaagaat 1898
 DB 16510 CTAAAGAAACGCTTCTTACTGCTCTTACATGTGATGGCATCTTACACAAAGAAATAT 16569
 OY 1899 tattgggggaaaaaacacaaagtgctgtgacatgctctgagaagaactgcagagacataagct 1958
 DB 16570 TATTGGGGGAAAAACATCAAGTGTGTACATATCTGAGAAATCGCAGACATATAGCT 16629
 OY 1959 gcaacccaacatccttttgaagaatcattccagagaacacttacttctgtgtagtt 2018
 DB 16630 GCGACCCAAAAATCTTTTGAAGAAATCATTTCCAGCAACCTTACTTCTGTGTGTGT 16689
 OY 2019 ttaattgtaaaaaaaagtttlaacagaagcacatgacatgatalgaagccctgcagga 2078
 DB 16690 TTAATGTPTAAAAAAGGTTTAAACAGAGACATGACATATGAAGCCTGCAGGA 16749
 OY 2079 ctggtgttttttggcaattcttcacgtgtgagactgtgtgcagaagaatgaagtagtg 2138
 DB 16750 CTGTCGCTTTTGTGCAATTCCTCACGTGGACTGTCTCCACAAAGATGAAGTGTGG 16809
 OY 2139 ttttaagaagttaaattacatatatttctcactaaagttaattatgcaaaagttt 2198
 DB 16810 TTTTAAAGAGTTAAATTCATATTTATTTATTTCTCACTAAATTTATTTATTCAAAAGTTT 16869
 OY 2199 tctgtgagaatgacaatgatttaattatgcttattgaaatgaagctgtgtctccaag 2258
 DB 16870 TCTTGTGAAGATGACAAATGTATATTCCTTTTNGAATTAACAGTGTGTCTTCCAGAG 16929
 OY 2259 tccagagacatgttataaagaacaatgaatcagtcgaag 2301
 DB 16930 TCCAGAGACATTTGTATTAAGACATGAATCATGACCGAAG 16972

RESULT 5
 AC003098/c 94752 bp DNA PRI 29-JAN-1998
 LOCUS Homo sapiens chromosome 17, clone HRPc905N1, complete sequence.
 DEFINITION AC003098
 ACCESSION AC003098
 VERSION AC003098.1 GI:2822155
 KEYWORDS HNG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 94752)
 AUTHORS Birren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P. and
 Lander,E.
 TITLE Homo sapiens chromosome 17, clone HRPc905N1
 JOURNAL Unpublished
 2 (bases 1 to 94752)
 REFERENCE AUTHORS
 Richardson,P., Fasman,K., McKernan,K., Munro,C., Nusbaum,C.,
 Birren,B., Lander,E., Baldwin,J., Baran,N., Cantu,C., Chang,A.,
 Cooke,P., Daly,M.J., Devon,K., Dewar,K., Durelle,B., Forrest,C.,
 Gage,D., Gensheimer,S., Geraldery,K., Gilmarlin,T., Hages,B.,
 Halphen,I., Harris,K., Howland,J.C., Huang,J., Hui,L., Jacotol,L.,
 Kirby,A., Lane,M., Mackenzie,J., Marquis,N., McDermott,J.,
 Molla,M., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A.,
 Peterson,K., Roberts,D., Rollins,G., Sarnaik,A., Shiu,P., Shyam,R.,
 Stikwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L.,
 Zentgraf,I., and Zody,M.
 TITLE Direct Submission
 JOURNAL Submitted (10-NOV-1997) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 94752)
 AUTHORS Birren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P.,
 Lander,E., Allen,N., Baker,J., Baldwin,J., Baran,N., Beckery,R.,

Boutwell, C., Byrne, S., Cantu, C., Castle, A., Cooke, P., Daly, M.J.,
 Depierre, E., Devon, K., Dewar, K., Durelle, B., Etemadi, S.,
 Ferreira, P., Forrest, C., Gage, D., Gardyna, S., Gensheimer, S.,
 Gierdys, K., Gilmartin, T., Gray, D., Hagos, B., Harris, K.,
 Horton, L., Howland, J.C., Hu, L., Jacotot, L., Linton, L.,
 Mackenzie, J., Margulis, N., McEwan, P., McGurk, A., Meldrum, J.,
 Molla, M., Morris, W., Morrow, J., Nachman, A., Naylor, J., O'Connor, T.,
 Rollins, B., Peterson, K., Ranganath, S., Riley, R., Roberts, D.,
 Rollins, G., Rossello, R., Roy, A., Shyam, R., Soohoo, S., Stillwell, J.,
 Stone, C., Strickland, C., Sydney, K., Tang, L., Vassiliev, H., Vo, A.,
 Wagner, A., Wheeler, J., Wu, Y., Ye, W.J., Zemtseva, I., Zhao, J. and
 Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (29-JAN-1998) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Jan 29, 1998 this sequence version replaced gi:2655504.
 The Staden databases, finishing information, and all
 chromatographic files used in the assembly of this clone are
 available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. &
 Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

Location/Qualifiers

source
 1.94752
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HRC905N1"
 /clone_lib="Peter de Jong/ human PAC library"
 /map="17"
 /chromosome="17"
 86..439
 repeat_region /rpt_family="LINE2"
 complement(972..1262)
 /rpt_family="AluSx"
 repeat_region /rpt_family="AluSx"
 complement(11884..11912)
 /rpt_family="CA)n"
 1912..1959
 repeat_region /rpt_family="CA)n"
 2633..2681
 repeat_region /rpt_family="MIR"
 2701..3002
 repeat_region /rpt_family="AluSx"
 complement(3874..4008)
 /rpt_family="LINE2"
 4513..4543
 repeat_region /rpt_family="AT-rich"
 complement(5202..5237)
 /rpt_family="CA)n"
 repeat_region /rpt_family="MIR"
 complement(5310..5406)
 /rpt_family="MIR"
 repeat_region /rpt_family="MIR"
 complement(5829..5868)
 /rpt_family="GC-rich"
 6485..6592
 repeat_region /rpt_family="MIR"
 complement(6973..7029)
 /rpt_family="MIR"
 repeat_region /rpt_family="MIR"
 complement(10402..10426)
 /rpt_family="GC-rich"
 repeat_region /rpt_family="MIR"
 complement(11421..11496)
 /rpt_family="LINE2"
 repeat_region /rpt_family="AluSx"
 complement(11505..11804)
 11811..11879
 repeat_region /rpt_family="AT-rich"
 11880..12093
 repeat_region /rpt_family="Aluub"
 complement(12104..12476)
 /rpt_family="LINE2"
 12478..12773
 repeat_region /rpt_family="AluSx"
 complement(12775..12917)
 /rpt_family="LINE2"
 repeat_region

repeat_region complement(13088..13510)
 /rpt_family="MLTIC"
 repeat_region complement(13544..13847)
 /rpt_family="AluSx"
 repeat_region complement(14110..14776)
 /rpt_family="LINE1"
 repeat_region complement(14477..15077)
 /rpt_family="AluSx"
 repeat_region complement(15079..15282)
 /rpt_family="LINE1"
 15954..16024
 repeat_region /rpt_family="AT-rich"
 18031..18178
 repeat_region /rpt_family="MIR"
 20398..20769
 repeat_region /rpt_family="LINE2"
 20893..21190
 repeat_region /rpt_family="AluSg"
 complement(21275..21473)
 /rpt_family="AluSg"
 repeat_region complement(21485..21556)
 /rpt_family="CA)n"
 21567..21906
 repeat_region /rpt_family="MER2"
 21907..21952
 repeat_region /rpt_family="MLT2CB"
 21974..22115
 repeat_region /rpt_family="MIR"
 22238..22540
 repeat_region /rpt_family="AluSx"
 22577..22633
 repeat_region /rpt_family="LINE2"
 complement(23190..23270)
 /rpt_family="LINE2"
 24196..24324
 repeat_region /rpt_family="FLAM_A"
 24329..24643
 repeat_region /rpt_family="AluSp"
 25542..25698
 repeat_region /rpt_family="LINE2"
 25805..26001
 repeat_region /rpt_family="MIR"
 26060..26174
 repeat_region /rpt_family="LINE2"
 26343..26549
 repeat_region /rpt_family="AluSx"
 complement(27491..27783)
 /rpt_family="AluSg"
 27862..27906
 repeat_region /rpt_family="LINE2"
 28068..28114
 repeat_region /rpt_family="MIR"
 complement(28329..28360)
 /rpt_family="AT-rich"
 28371..28652
 repeat_region /rpt_family="AluSg"
 complement(29498..29604)
 /rpt_family="GC-rich"
 30074..30181
 repeat_region /rpt_family="MIR"
 complement(30181..30249)
 /rpt_family="LINE2"
 30444..30517
 repeat_region /rpt_family="MIR"
 complement(31852..32135)
 /rpt_family="AluSg"
 repeat_region complement(32675..33003)
 /rpt_family="MLT1A1"
 complement(33042..33082)
 /rpt_family="LINE2"
 33363..33459
 repeat_region /rpt_family="MER5B"
 complement(35475..35524)
 repeat_region

```
/rpt_family="MLT2D"
35495..35591
/rpt_family="(GA)n"
complement(35561..35673)
repeat_region
/rpt_family="MLT2D"
complement(35678..35935)
repeat_region
/rpt_family="MER6B"
complement(35936..36233)
repeat_region
/rpt_family="Alusg"
complement(36236..36449)
repeat_region
/rpt_family="MER6B"
complement(36450..36867)
repeat_region
/rpt_family="MLT2D"
complement(36876..37460)
repeat_region
/rpt_family="L1MD2"
complement(37477..37777)
repeat_region
/rpt_family="Alusx"
complement(37804..37948)
repeat_region
/rpt_family="L1"
37948..38130
repeat_region
/rpt_family="L1MD1"
38175..38253
/rpt_family="LINE2"

Query Match      89.0%; Score 2047.8; DB 85; Length 94752;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2061; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```
|||||
Db 5680 TGAAGGTCCACAGGGGAGGAGGAATTTGAGATCTACAGACTGAGCCAGCC 5621
QY 939 ccgcctctggggccgcctaccttgcttgctccaccttcagagagagagaaatggaaga 998
Db 5620 CCCCTCTGGGGCCGCCTACCTTGTGCTGCCACTTCAGAGAGGACAGAAATGGAAGA 5561
QY 999 ttctcaacggccctggggttttaaggagagcggtgttggaatggaagttccagagactgt 1058
Db 5560 TTTTCACCGCCCTGGGGTWTAAAGGAGCGGTGTGGAGTGGAGAAAGTCCAGSGACTGCT 5501
QY 1059 taagaaagtgtgaataagatcccccttgcaacctgcgtccctcaagaagaagcctggaggt 1118
Db 5500 TAAGAAAGTTGAGTAAGATATCCCCCTTGACACCTGCGCATCAGAAAGCCTGAGGCT 5441
QY 1119 gcccaagacaaagactggggggaactgtagatggtttcctagctcgtcgtccact 1178
Db 5440 GCCCAGAGCACAAGACTGGGGGCAACTGTAGATGTGTGTTTACGCTTGCTCCACT 5381
QY 1179 aacttgctgtgaaccttgaaactacacaacttcctctggagacctcaattccacttgt 1238
Db 5380 AACTTGCTGTGAACCTTGAACTACACAAATTCCTTCGGACCTCAATTCCTACTTGT 5321
QY 1239 aaaaatgaggttggaaggtggaataagatctcgagagaaactatggtacatagtatccaa 1298
Db 5320 AAATGAGGGGTGAGGTGGGAATAGATCTCCAGAGAGACTATTTGGCATATGATTCAGAG 5261
QY 1299 actccaagtcccttttaatgagagagagtgtgagagagagagagagagagaaatgaa 1358
Db 5260 ACTCCAGTGCCTTTTAAATGGGAGAGAGTGAAGAGAGAGAGAAAGAGATGAATGA 5201
QY 1359 tgcagttgcatlgaatcagagtgcaaggtgacatccagaattcagaagttgttgatctct 1418
Db 5200 TGCAGTTGCAATTATTCAGTGCCTCAAGTCACTTCCAGATTTAGAACTTGTGATGCTCT 5141
QY 1419 tctgacagccaaagatgaaataaacaagaaaaaaaglaaagatctatltatgtct 1478
Db 5140 TCTGACAGCCAAAGATGAAATAACAACAAGAAAAAAAGTAAAGTATTAATTAATGACT 5081
QY 1479 gaatatlttcaggctgaacaactcctcggaagaagatgacatgctcttccagactgcttc 1538
Db 5080 GACATATTTATGSGCTGACAAACCTCTGGAAGAGATATCTGCTCCAGCTGACTTCC 5021
QY 1539 ccgagttgttgctacctcaaccctccatctcaagaataacatccatctctgagta 1598
Db 5020 CCGGATTTGGCTACTCTCACCCCTCCATCTTAAGAATAATACATCACTTGGGGTA 4961
QY 1599 gaaagagagaggttcgagaggtgtgtgagagagatagaatcaatccgcaccaactccc 1658
Db 4960 GAAAGAGAGAGCGCTCGAGGCTGTGAGGAGGATGAATAATCACATCCGCCCAACTTCC 4901
QY 1659 aaagagagatccctcccccgagcccatagccatgltttaaagttacacttccgaagaa 1718
Db 4900 AAAGAGCAGATCCCTCCCGAGCCCATAGCCATGTGTTTAAAGTACCTTCCGAAGAGAA 4841
QY 1719 gtgaaggtltcaagagacactggtccttcagagcccgagagagagccatcacaaactaca 1778
Db 4840 GTGAAGGTTCAAGGAGACACTGGCTTGCAAGGCCCGAGGAGAGCAGCATCAAACTCACA 4781
QY 1779 gaccagacatccctlttgagacacgctcttgcaccactcaacgagacacattctgctc 1838
Db 4780 GACCAGACATCCCTTTTGAGACACCGCCTTGCCACACATCAGGACGACATTTCTGTC 4721
QY 1839 ctgaagaacagcttcttactagctcttaactgtagatgagatattcttaactaaagaatat 1898
Db 4720 CTGAAGAAACGCTTCTTACTGCTTTCATCATGTGATGCAATATCTTAATAAAGAAAT 4661
QY 1899 tattgaggagaaaaactacaagtgctgtacatagtctgaaanaactcgagagacataagct 1958
Db 4660 TATTTGGGAGAAAAACTACAAGTGTCTGATACATATGCTGAGAAACTGAGAGCATTAAT 4601
QY 1959 gccacccaaaatctcttttgaaaatcaattccaagacacttacttctgtgtgagttt 2018
```



```

Db 113251 TCTGACAGCCAAAGATGAAACAAAGAAAAAAGTAAAGATCTATTATGGCT 113310
QY 1479 gacatattacggctgacaaacccctcggaagaagctatgctctccagcgtgcttcc 1538
Db 113311 GACATATTTCAGGCTGCAAAACCTCGGAAGAGATATGCTTCCAGCCGCTTCC 113370
QY 1539 ccgagatcttggcttactccaccctccatctcaagaataacatccatcttggtgta 1598
Db 113371 CCGGATTTGGGCTACTCCACCCTCCATCTCAAGAAATTAATCATCATTCATGGGGTA 113430
QY 1599 gaaaagagaagggtccgaggtggtggaaggaatagaatcacatccgcccaacttccc 1658
Db 113431 GAAAAGAGAGAGGGTCCGAGGGTGGTGGAGGATAGAAATCATCATCCGCCCAACTTCCC 113490
QY 1659 aaaaagagagatccctcccccagccatgccaagtgttaagtcacatcccttcggaaga 1718
Db 113491 AAAGAGAGAGATCCCTCCCGACCCATACCATGCTTTAAAGTCACTTCCGAAGAGAA 113550
QY 1719 gtgaaggttcaagagacatgcttgcagcccgagggagagacatcaacaactcaca 1778
Db 113551 GTGAAGGCTTCAAGAGACATGGGCTTGCAGGCCCGAGGGAGCAGCATCAAAACTCACA 113610
1779 gaccagacatccctcttgagacacgcgctctgcccacactcacgagacatcttcgc 1838
Db 113611 GACCAGACATCCCTTTTGAGACACGCGCTTCCGCCACACACACGACATTTCTGC 113670
QY 1839 ctgaaaagagctctctactgctctacatgtagatgcatcttactcaaaagaatg 1898
Db 113671 CTAGAAAACAGCTTCTTACTGCTTTCATGATGATGAGCATATCTTACTAAAGAAATAT 113730
QY 1899 tatctgggggaaaaaactaacaagtctgtacatctgtagaagaactgcagagacataatg 1958
Db 113731 TATGGGGGAAAAACTACAGAGTCTGTACATATGTGAGAACTGCAGAGCATATAATA-CT 113789
QY 1959 gccaccccaaaaactctttaaataatcattccagacaaactctacttctgtgtagtt 2018
Db 113790 GCCACCCAAAAATCTTTTAAATCATTTCCAGACAACTCTTACTTCTGTGTGATTT 113849
QY 2019 ttaattgttaaaaaaaagttttaaacaagaagacatgacatgaaagcctgcagga 2078
Db 113850 TTAATTGTTAAAAAAGTTTAAACGAAGACATGACATATGAAAGCCTGCAGGA 113909
QY 2079 ctggtcgttcttggcaatcttccagctggaactgtcccaagaatgaagtgtg 2138
Db 113910 CTGTCCTTTTGGCAATCTTCCACGTGGGAGCTTGCACAAAGATGAAGTGTG 113969
QY 2139 ttcttaagaggttaagtacatattatcttccacttaagtattatgcaaaagttt 2198
Db 113970 TTTTAAAGGTTTAAGTTTAAATTTTCTTCTACTTAAGTTTATGCAAAAGTTT 114029
2199 tcttgtagaagaatgacaatgttaataatgcttataatgaatcaagctgttcttccagag 2258
Db 114030 TCTGTGAGAGATGACAAATGTTAATTTGCTTATGATTAACAGTCTGTTCACAG 114089
QY 2259 tccaagagacatgttaataagaacaatgaatcatgaccgaag 2301
Db 114090 TCCAGAGACATTTGTTAATAAGCAATGATCATGACCGAAG 114132

```

```

RESULT 7
AC073954/c AC073954 177744 bp DNA HTG 17-NOV-2000
LOCUS Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
DEFINITION SEQUENCE, 34 unordered pieces.
ACCESSION AC073954
VERSION AC073954.3 GI:10998993
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 177744)
AUTHORS Waterston,R.H.

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 177744)
Waterston,R.H.
Direct Submission
Submitted (07-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Oct 25, 2000 this sequence version replaced gi:1186649.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H_NH0147C10
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.99019
Consensus quality: 153679 bases at least Q40
Consensus quality: 163329 bases at least Q30
Consensus quality: 166780 bases at least Q20
Insert size: 18500; agarose-fp
Insert size: 183559; sum-of-contigs
Quality coverage: 3.29 in Q20 bases; agarose-fp
Quality coverage: 3.43 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1247:	contig of 1247 bp in length
*	1248	1347: gap of unknown length
*	1348	2762: contig of 1415 bp in length
*	2763	2862: gap of unknown length
*	2863	4773: contig of 1911 bp in length
*	4774	4873: gap of unknown length
*	4874	6326: contig of 1453 bp in length
*	6327	6426: gap of unknown length
*	6427	8873: contig of 2447 bp in length
*	8874	8973: gap of unknown length
*	8974	10783: contig of 1810 bp in length
*	10784	10883: gap of unknown length
*	10884	12763: contig of 1880 bp in length
*	12764	12863: gap of unknown length
*	12864	15381: contig of 2518 bp in length
*	15382	15481: gap of unknown length
*	15482	17365: contig of 1884 bp in length
*	17366	17465: gap of unknown length
*	17466	19992: contig of 2527 bp in length
*	19992	20092: gap of unknown length
*	20093	23512: contig of 3420 bp in length
*	23513	23612: gap of unknown length
*	23613	25699: contig of 2087 bp in length
*	25700	25799: gap of unknown length
*	25800	28333: contig of 2534 bp in length
*	28334	28433: gap of unknown length
*	28434	32027: contig of 3594 bp in length
*	32028	32127: gap of unknown length
*	32128	36115: contig of 3988 bp in length
*	36116	36215: gap of unknown length
*	36216	39882: contig of 3667 bp in length
*	39883	39982: gap of unknown length
*	39983	44009: contig of 4027 bp in length
*	44010	44109: gap of unknown length
*	44110	47829: contig of 3720 bp in length

*	47830	47929	gap of unknown length
*	47930	53768	contig of 5839 bp in length
*	53769	53868	gap of unknown length
*	53869	60215	contig of 6347 bp in length
*	60216	60315	gap of unknown length
*	60316	66477	contig of 6162 bp in length
*	66478	66577	gap of unknown length
*	66578	71476	contig of 4899 bp in length
*	71477	71576	gap of unknown length
*	71577	78883	contig of 7307 bp in length
*	78884	85852	gap of unknown length
*	78884	85852	contig of 6969 bp in length
*	85953	86052	gap of unknown length
*	86053	93510	contig of 7458 bp in length
*	93511	93610	gap of unknown length
*	93611	100718	contig of 7108 bp in length
*	100719	100818	gap of unknown length
*	100819	109795	contig of 8977 bp in length
*	109896	109895	gap of unknown length
*	109896	117790	contig of 7895 bp in length
*	117791	117890	gap of unknown length
*	117891	126044	contig of 8154 bp in length
*	126045	126144	gap of unknown length
*	126145	134442	contig of 8298 bp in length
*	134443	134542	gap of unknown length
*	134543	145663	contig of 11121 bp in length
*	145664	145789	contig of 11426 bp in length
*	145790	157189	contig of 11426 bp in length
*	157190	157289	gap of unknown length
*	157290	166822	contig of 9533 bp in length
*	166923	166922	gap of unknown length
*	166923	177744	contig of 10822 bp in length.

UNES	Location/Qualifiers
source	1..177744
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="17"
	/clone="RP11-147C10"
misc_feature	1..1247
misc_feature	/note="assembly_name:Contig26"
misc_feature	1348..2762
misc_feature	/note="assembly_name:Contig31
misc_feature	2863..4773
misc_feature	/note="assembly_name:Contig34"
misc_feature	4874..6326
misc_feature	/note="assembly_name:Contig35
misc_feature	6427..8873
misc_feature	/note="assembly_name:Contig36"
misc_feature	8974..10783
misc_feature	/note="assembly_name:Contig37
misc_feature	10884..12763
misc_feature	/note="assembly_name:Contig38
misc_feature	12864..15381
misc_feature	/note="assembly_name:Contig39
misc_feature	15482..17465
misc_feature	/note="assembly_name:Contig40"
misc_feature	17466..19992
misc_feature	/note="assembly_name:Contig41
misc_feature	20093..22312
misc_feature	/note="assembly_name:Contig42
misc_feature	22613..22609
misc_feature	/note="assembly_name:Contig43
misc_feature	22800..28333
misc_feature	/note="assembly_name:Contig44
misc_feature	28434..32027
misc_feature	/note="assembly_name:Contig45
misc_feature	32128..36115
misc_feature	/note="assembly_name:Contig46
misc_feature	36216..33982
misc_feature	/note="assembly_name:Contig47
misc_feature	33983..44009
misc_feature	/note="assembly_name:Contig48
misc_feature	44110..47829

```
misc_feature /note="assembly_name:Contig49"
47930. .53768
/note="assembly_name:Contig50"
misc_feature 53869. .60215
/note="assembly_name:Contig51"
misc_feature 60316. .66477
/note="assembly_name:Contig52"
misc_feature 66578. .71476
/note="assembly_name:Contig53"
misc_feature 71577. .78883
/note="assembly_name:Contig54"
misc_feature 78984. .85952
/note="assembly_name:Contig55
clone_end:r7
vector_side:left"
misc_feature 86053. .93510
/note="assembly_name:Contig56"
misc_feature 93611. .100718
/note="assembly_name:Contig57"
misc_feature 100819. .109795
/note="assembly_name:Contig58"
misc_feature 109896. .117790
/note="assembly_name:Contig59"
misc_feature 117891. .12604
/note="assembly_name:Contig60"
misc_feature 126145. .134442
/note="assembly_name:Contig61"
misc_feature 134543. .145663
/note="assembly_name:Contig62"
misc_feature 145764. .157189
/note="assembly_name:Contig63"
misc_feature 157290. .16682
/note="assembly_name:Contig64"
misc_feature 166923. .177744
/note="assembly_name:Contig65"

BASE COUNT 44147 a 43012 c 43431 g 43839 t 3315 others
ORIGIN
```

Query Match	46.3%	Score 1066.4	DB 75	Length 177744
Best Local Similarity	97.0%	Pred. No. 5.6e-195		
Matches 1129	Conservative	0	Mismatches 31	Indels 4
				Gaps 4
OY 1142	aactgatagatggttctt-agtctggtctgccc-aaactgtctgttaacctgaa	1199		
Db 93369	AGCTGGAGATGTTGTCTGAGACTTCCTGCGCTGCGACCTTACGTGGGAAATCCCTGAA	93310		
OY 1200	ctaacacattctcttcgagactcaattcca-cttggtaaatagagttgagttgag	1258		
Db 93309	CTACACAAAGTCGTGGCGGGACCTCATAGTCCATCTGTGTAATAATGATGTTGAGAGTTGGG	93250		
OY 1259	aatgagatctcgagagactattgcatatgattccaaagactcagttgcttltgaatg	1318		
Db 93249	AATGGGATTCGAGGGAGACTATTGGCAATAGTCGCAAGGACTCCAGTGGCTTTTGAAAG	93190		
OY 1319	ggcagaggttaagagagagagagagagagagaaatgaatgacgttgcattatcatt	1378		
Db 93189	GGCAGAGCTAGAGAGAGAGAGAGAGAGAGAGAGATGATGCAATGCTTATTCATTCAT	93150		
OY 1379	gccaaagtcactctccagaatt-cagaattgtgagtctctctcttgcagaccaaagatgaa	1437		
Db 93129	GCCAAAGTCACTTCCAGAAATTTCAGAGTTTGATGCTCTCTTCGACAGCCAAAGATGAA	93070		
OY 1438	aaacaaacacaaaaaaaagaatgaagatcatctttaaagctgcacataataaagcgtgaca	1497		
Db 93069	AAACAAACGAAAAAAGTAAGAGTATTATTATGCGTGCACATATTACGCGTGACA	93010		
OY 1498	aaactcttggaagaagcatatgctgtcttccagagcttgccttcccgatgtttgactactc	1557		
Db 93009	AACTCCTGGGAAGGATGATGCTGCTCCACAGCTGGCTTCCCGAGTTTGCTATCCAC	92950		
OY 1558	caacctccatctcagaagaatacaacatcatcatattgagtgagtaaaagagaggtccgag	1617		

Db	92349	CACCCTCCATCTCCAAAGAAATTAACATCATTCATTTGGGGTGTGAAAGAGAGAGGCTCCGAG	92890
Qy	1618	ggttggtagggagtagaatacacatccgcccacaactcccaagaagcaatccctccc	1677
Db	92889	GGTGTGTGGGGGATGTGAATACATCCGCCCAACTTCCCAAGAGACAGCATCCCTCC	92830
Qy	1678	ccgaccacatagcagatgtttaagaagcaaccttcccgagaagaagtgaagaagttcagaagca	1737
Db	92829	CCGACCCATAGCCATGTTTAAAGTCACCTTCCGAAGAAAGTGAAGGTTCAAGGACAC	92770
Qy	1738	tggccttgcagcgcgaggaagcagacalcaacaactcagacagacagacatccctttg	1797
Db	92769	TGGCTTTGCAGGCGCCAGGAGGACAGCCATCAACAACTCACAGACACACATCCCTTTG	92710
Qy	1798	agacacccgctcttgcgcaacactcaagagacacatcttctgcctagaagaacagctcttac	1857
Db	92709	AGACACGCGCTTGTGCCACACACTCAGGACACATTTGTGCTGTGAAGAACAGCTTCTTAC	92650
Qy	1858	tgcctctacaatgtgatgtgcatactcttaacaataaagaatattatcttggtgggaaaaactaa	1917
Db	92649	TGCTCTTACATGTGATGGGATATCTTTCACATAAAGAAATATATTGGGGGAAAAAACATCA	92590
Qy	1918	agtcgctgacatagtctgagaacatgcgcagagacataatagctgcaccccaaatctttt	1977
Db	92589	AGTGTCTACATATGTGTGAGAAACATGCAGAGCATATATCTGCCACATCAAAATCTTTT	92530
Qy	1978	gaaatacatttcagagaacacctctactcttctgtgtagtttcttaattgtttaaaaaaaa	2037
Db	92529	GAAATTCATTTCAGAGCAACCTTACCTTTCGTGTGTAATTTTAAATGTTAAAAA	92470
Qy	2038	agtttbaacagaagaacacatgacataatgaagcctgcgcagacatgctgtcttttggcaa	2097
Db	92469	AGTTTAAACAGAGACATGACATATGAAAGCCGACGAGCATGTGCTTTTGTGGCAA	92410
Qy	2098	ttcttcacagtggagactgttccaaagaatgaagtagtggtttcttaagaagcttaagta	2157
Db	92409	TTCCTCCAGCTGGGACTGTTCACACAGAAATGAAATGTGTGTTTAAAGAGTTAAATTA	92350
Qy	2158	catacttatcttccacttaagtaatttatagaaaagcttctctctgtagagaatgaacaat	2217
Db	92349	CATTTTATTTTCTCACTTAAGTTATTTATGCAAAAGTTTCTTGTAGAGAAATGACAAAT	92290
Qy	2218	gtaataatgtccttaagaatgaagtcgttcttcgagagatccagagaatgttaata	2277
Db	92289	GTTAATATTCCTTANTGAATTAACAGTCTGTTCTTCAGAGTCCAGAGACATGTGTTAATA	92230
Qy	2278	aagacaatgaatcatgaccgaag	2301
Db	92229	AAGCAATGATCATGACCGGAAG	92206
LOCUS	AF326742	642 bp	mRNA
DEFINITION	Cercopithecus aethiops sclerostin mRNA, complete cds.		28-FEB-2001
VERSION	AF326742		
ACCESSION	AF326742.1	GI:13161028	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			

AUTHORS		TITLE		JOURNAL	
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepert, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beignton, P. and Mulligan, J.T.		Direct Submission		Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th st. SE, Bothell, WA 98021, USA	
FEATURES		Location/Qualifiers			
Source		1..642		/organism="Ceropithecus aethiops"	
CDS		/db_xref="taxon:9534"			
		1..642			
		/codon_start=1			
		/product="sclerostin"			
		/protein_id="AAK13457.1"			
		/db_xref="GI:13161029"			
		/translation="MOPLALCIYCLIVLVAHAFVVEGGMQAFKNDAPETIPELCEYV EPPPELENNKTNRAENGRPRHPFETKVKSEYSECRETAFRTYTDGCRSAKRYVE LVSGGCGGPARLLPNAIGRGMWRPSPDPRCLIPDRKORVOLCPGGAARAKRYVE LVASCKRKLTFEHNOSLEKDFGEARLPQGRKPRPARAKAMQAELENNY"			
BASE COUNT		113 a 228 c 210 g 91 t			
ORIGIN					
Query Match 26.5%; Score 610; DB 89; Length 642;					
Best Local Similarity 96.9%; Pred. No. 4.1e-107;					
Matches 622; Conservative 0; Mismatches 20; Indels 0; Gaps 0;					
QY	48 atgcagctccacatgacgctgtgtgttcacatctgcgcctggtgtacacacagcttcctgtgta	107			
Db	1 ATGCAGCTCCACACTGCGCCCTGTCTCTTCTCTCTGCTGTGTACACGACGCTTCCGCTGA	60			
QY	108 gtgtgagggcccaaggggtgtgcaagggttcacaaatgtatgcacagaaatcatcccgagagac	167			
Db	61 GTTGAGGGCCCGAGGGGTGTCAGGCTTTCACGAATGATGCCAGGAAATCATCCGAGACTTC	120			
QY	168 ggaagatfaccccgagcctccacccgagcctgtgagaacaaacagacatgtaaaccggtcgagaa	227			
Db	121 GGAGAGTACCCCGACGCTCCACCGGAGCTGGAGAACAAACAGACCATGTAACCGGGCGGAG	180			
QY	228 aagcgaaggcgagctcccccacaccccccttttgagaccacaaagcgtgtccgagtaacagctgc	287			
Db	181 AATGGAGGGGGGCTTCCCAACACCCCTTTGAGAACCAAGACGCTGTCCGAGTACAGCTGC	240			
QY	288 cgcagactgtgaattcacccgctacgtgtgacgcgtgtggccgtgtgcgacgacgcgaagccgctg	347			
Db	241 CGAAGCTGTGACTTACACCCGCTACGAGACCGATGGGCGCTCCGACGCGCAGACCAAGTC	300			
QY	348 accgagctgtgtgtgtccctccgacgctgtgcgcgcgcgcgtgtgtccacaaagccatctgcgc	407			
Db	301 ACCGAGTGTGGTGTCTCCGCGCAGTGTGCGGCCCGGCAAGCGCTGTGCCAACGCCATTCGCG	360			
QY	408 cgcggcgaagtgtgtgtgacactagtgtggcccgacttcgcgtgtgcatcccgacgctacgacgc	467			
Db	361 CGCGCGCAAGTGTGTGGCCCGAGATGGGCGCCGACCTTCCTCATCTCCCGACCGCTACCG	420			
QY	468 ggcgcagcgcgtgcagcgtgtgtgtcccggtgtgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgc	527			
Db	421 GCGCAGCGCTGTGCAGCTGCTGTCTCCGCGTGTGTCGCGCGCGCGCGCGCGCGCGCGCGCG	480			
QY	528 ctgtgtgctctgtgtgcaagtgtgcaagcgcgtccacacacgcttcgagctcgaag	587			
Db	481 CTGTGTGCTCTGTGTCAAGTGTCAAGGCCCTTCACACCATCTTCACAAACAGTGTGAGCTCAAG	540			
QY	588 gacttcggagacgaagcgcgtgtgc	647			
Db	541 GACTTCGCTGTCCGAGCGCGCTGTCCGCGCAGAAAGGCGCGGAAGCCGCGCGCGCGCGCGCG	600			
QY	648 agcgcgaagcccaacgaagc	689			
Db	601 GGGCGCAAAAGCAATCAAGCGCGCGCTGTGGAAGACCGCTCTGTG 642				

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
314	CCGGAGTGGCCAAACCGCGTACCGAATTGGTGTCTCGTGGCGCCAGTGGCGCCGCGCGCT			373			
389	gctgtcccaacgcatctcgccgagcaagtgttgagaaactaagtggccggaactccgtg			448			
374	GTGTGCCCAAGCGCATGTGGGCGGTGAATGGTGGCGGCCCAAGAGACCGGACTTCCGTG			433			
449	catcccgacgcgtacacgcgcgcagcgcgttcagctgctgtgccggtgtgagagcgc			508			
434	CATCCGAGATCGCTACCGCGCGCACCGGGGTGACCTCTGTGGCCCGCGCGCGCGCGC			493			
509	ggcgcgagccgaagagtgagcgtgtgtgctctgtgtgaatgaacgcgtaccgcgttca			568			
494	GGCTCGCGCCAAAGTGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT			553			
569	caaccagtcgagagctcaaggaactctggagaccgaagccgcctcgccgcgaagagccgaa			628			
554	CAACCAATCGAGCTTCAGAGACTTGTGGACCTGAGACCCGCGCCGACAGAGGTCCGA			613			
629	ggcgcgagcccgccgcccgcgcgcgcgaagcaaacgacagcgcgcgtgagaaagccacta			688			
614	GGCGGCGCCCGCGCGCGCGCGGAGCCAAAGCCAACAGCGGAGCTGAGAGAGCCCTACTA			673			
Qy	689 g	689					
Db	674 G	674					
RESULT	10						
AC073954							
LOCUS	AC073954						
DEFINITION	AC073954	177744 bp	DNA	HTG	17-NOV-2000		
ACCESSION	AC073954						
VERSION	AC073954.3	GI:10988993					
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.						
ORGANISM	human.						
REFERENCE	Human sapiens						
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
JOURNAL	1 (bases 1 to 177744)						
REFERENCE	Waterston, R.H.						
AUTHORS	The sequence of Homo sapiens clone						
TITLE	Unpublished						
JOURNAL	2 (bases 1 to 177744)						
REFERENCE	Waterston, R.H.						
AUTHORS	Direct Submission						
TITLE	Submitted (07-JUL-2000) Genome Sequencing Center, Washington						
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis,						
REFERENCE	MO 63108, USA						
AUTHORS	On Oct 25, 2000 this sequence version replaced gi:1986849.						
TITLE	COMMENT						
JOURNAL							
REFERENCE	Genome Center						
AUTHORS	Center: Washington University Genome Sequencing Center						
TITLE	Center code: WUGSC						
JOURNAL	Web site: http://genome.wustl.edu/gsc/Index.shtml						
REFERENCE	Project Information						
AUTHORS	Center project name: H.NR0147C10						
TITLE	Summary Statistics						
JOURNAL	Sequencing vector: M13; 100%						
REFERENCE	Sequencing vector: plasmid; 0%						
AUTHORS	Chemistry: Dye-primer ET; 100% of reads						
TITLE	Chemistry: Dye-terminator Big Dye; 0% of reads						
JOURNAL	Assembly program: Phrap; version 0.990319						
REFERENCE	Consensus quality: 155679 bases at least Q40						
AUTHORS	Consensus quality: 163329 bases at least Q30						
TITLE	Consensus quality: 166780 bases at least Q20						
JOURNAL	Insert size: 185000; agarose-fp						
REFERENCE	Insert size: 183559; sum-of-contigs						
AUTHORS	Quality coverage: 3.29 in Q20 bases; agarose-fp						
TITLE	Quality coverage: 3.43 in Q20 bases; sum-of-contigs						
JOURNAL	* NOTE: This is a working draft sequence. It currently						

* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1247: contig of 1247 bp in length
* 1248 1347: gap of unknown length
* 1348 2762: contig of 1415 bp in length
* 2763 2862: gap of unknown length
* 2863 4773: contig of 1911 bp in length
* 4774 4873: gap of unknown length
* 4874 6326: contig of 1453 bp in length
* 6327 6426: gap of unknown length
* 6427 8873: contig of 2447 bp in length
* 8874 8974 10783: gap of unknown length
* 8974 10783: contig of 1810 bp in length
* 10784 10883: gap of unknown length
* 10884 12763: contig of 1880 bp in length
* 12764 12863: gap of unknown length
* 12864 15381: contig of 2518 bp in length
* 15382 15481: gap of unknown length
* 15482 17365: contig of 1884 bp in length
* 17366 17465: gap of unknown length
* 17466 19992: contig of 2527 bp in length
* 19993 20092: gap of unknown length
* 20093 23512: contig of 3420 bp in length
* 23513 23612: gap of unknown length
* 23613 25699: contig of 2087 bp in length
* 25700 28333: contig of 2534 bp in length
* 28334 28433: gap of unknown length
* 28434 32027: contig of 3594 bp in length
* 32028 32127: gap of unknown length
* 32128 36115: contig of 3988 bp in length
* 36116 36215: gap of unknown length
* 36216 39882: contig of 3667 bp in length
* 39883 39983: gap of unknown length
* 39983 44009: contig of 4027 bp in length
* 44010 47829: gap of unknown length
* 47829 47930: contig of 3720 bp in length
* 47930 53768: gap of unknown length
* 53769 53868: contig of 5839 bp in length
* 53869 60215: gap of unknown length
* 60216 60315: contig of 6347 bp in length
* 60316 66477: gap of unknown length
* 66478 66577: contig of 6162 bp in length
* 66578 71476: gap of unknown length
* 71477 71576: gap of unknown length
* 71577 78883: contig of 7307 bp in length
* 78884 85952: gap of unknown length
* 85953 86052: contig of 6969 bp in length
* 86053 93510: gap of unknown length
* 93511 93610: contig of 7458 bp in length
* 93611 100718: gap of unknown length
* 100719 100818: contig of 7108 bp in length
* 100819 109795: contig of 8977 bp in length
* 109796 109895: gap of unknown length
* 109896 117790: contig of 7895 bp in length
* 117791 117890: gap of unknown length
* 117891 126044: contig of 8154 bp in length
* 126045 126144: gap of unknown length
* 126145 134442: contig of 8298 bp in length
* 134443 134542: gap of unknown length
* 134543 145663: contig of 11121 bp in length
* 145664 145763: gap of unknown length
* 145764 157189: contig of 11426 bp in length
* 157190 157289: gap of unknown length
* 157290 166822: contig of 9533 bp in length
* 166823 166922: gap of unknown length

FEATURES * 166923 177744: contig of 10822 bp in length.
source
1. 177744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-147C10"
1. 1247
/note="assembly_name:Contig26"
1348..2762
/note="assembly_name:Contig31"
2863..4773
/note="assembly_name:Contig34"
4874..6326
/note="assembly_name:Contig35"
6427..8873
/note="assembly_name:Contig36"
8974..10783
/note="assembly_name:Contig37"
10884..12763
/note="assembly_name:Contig38"
12864..15381
/note="assembly_name:Contig39"
15482..17365
/note="assembly_name:Contig40"
17466..19992
/note="assembly_name:Contig41"
20093..23512
/note="assembly_name:Contig42"
23613..25699
/note="assembly_name:Contig43"
25800..28333
/note="assembly_name:Contig44"
28434..32027
/note="assembly_name:Contig45"
32128..36115
/note="assembly_name:Contig46"
36216..39882
/note="assembly_name:Contig47"
39983..44009
/note="assembly_name:Contig48"
44110..47829
/note="assembly_name:Contig49"
47930..53768
/note="assembly_name:Contig50"
53869..60215
/note="assembly_name:Contig51"
60316..66477
/note="assembly_name:Contig52"
66578..71476
/note="assembly_name:Contig53"
71577..78883
/note="assembly_name:Contig54"
78984..85952
/note="assembly_name:Contig55"
vector_side:left"
clone_end:17
86053..93510
/note="assembly_name:Contig56"
93611..100718
/note="assembly_name:Contig57"
100819..109795
/note="assembly_name:Contig58"
109896..117790
/note="assembly_name:Contig59"
117891..126044
/note="assembly_name:Contig60"
126145..134442
/note="assembly_name:Contig61"
134543..145663
/note="assembly_name:Contig62"
145764..157189
/note="assembly_name:Contig63"

[illegible]

Consensus quality: 181484 bases at least Q40
Consensus quality: 190159 bases at least Q30
Consensus quality: 193745 bases at least Q20
Insert size: 207000; agarose-fp
Insert size: 197427; sum-of-contrigs
Quality coverage: 3.8 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contrigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1091: contrig of 1091 bp in length
* 1092 1191: gap of 100 bp
* 1192 3065: contrig of 1874 bp in length
* 3066 3165: gap of 100 bp
* 3166 4405: contrig of 1140 bp in length
* 4406 4405: gap of 100 bp
* 4406 5983: contrig of 1578 bp in length
* 5984 6083: gap of 100 bp
* 6084 8322: contrig of 2239 bp in length
* 8323 8423: gap of 100 bp
* 8423 10098: contrig of 1676 bp in length
* 10099 10198: gap of 100 bp
* 10199 12592: contrig of 2394 bp in length
* 12593 12692: gap of 100 bp
* 12693 14684: contrig of 1992 bp in length
* 14685 14784: gap of 100 bp
* 14785 17163: contrig of 2379 bp in length
* 17164 17263: gap of 100 bp
* 17264 19900: contrig of 2637 bp in length
* 19901 20000: gap of 100 bp
* 20001 23054: contrig of 3054 bp in length
* 23055 23154: gap of 100 bp
* 23155 26011: contrig of 2857 bp in length
* 26012 26111: gap of 100 bp
* 26112 29412: contrig of 3301 bp in length
* 29413 29512: gap of 100 bp
* 29513 32991: contrig of 3479 bp in length
* 32992 33091: gap of 100 bp
* 33092 36913: contrig of 3822 bp in length
* 36914 37013: gap of 100 bp
* 37014 40880: contrig of 3867 bp in length
* 40881 40980: gap of 100 bp
* 40981 45015: contrig of 4035 bp in length
* 45016 45115: gap of 100 bp
* 45116 50321: contrig of 5206 bp in length
* 50322 50421: gap of 100 bp
* 50422 54284: contrig of 3863 bp in length
* 54285 54384: gap of 100 bp
* 54385 59430: contrig of 5046 bp in length
* 59431 59530: gap of 100 bp
* 59531 64373: contrig of 4843 bp in length
* 64374 64473: gap of 100 bp
* 64474 70503: contrig of 6030 bp in length
* 70504 70603: gap of 100 bp
* 70604 77829: contrig of 7226 bp in length
* 77830 77939: gap of 100 bp
* 77939 85976: contrig of 8047 bp in length
* 85977 86076: gap of 100 bp
* 86077 91694: contrig of 5618 bp in length
* 91695 91794: gap of 100 bp
* 91795 99208: contrig of 7414 bp in length
* 99209 99308: gap of 100 bp
* 99309 107493: contrig of 8191 bp in length
* 107500 107599: gap of 100 bp
* 107600 115731: contrig of 8132 bp in length
* 115732 115831: gap of 100 bp
* 115832 123297: contrig of 7466 bp in length

FEATURES
source

123298 123397: gap of 100 bp
* 123398 133979: contrig of 10862 bp in length
* 133980 134079: gap of 100 bp
* 134080 144631: contrig of 10552 bp in length
* 144632 144731: gap of 100 bp
* 144732 159118: contrig of 14387 bp in length
* 159119 159218: gap of 100 bp
* 159219 172437: contrig of 13219 bp in length
* 172438 172537: gap of 100 bp
* 172538 200727: contrig of 28190 bp in length.
location/Qualifiers

1. .200727
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-346P7"
/clone_11b="RPC1-23 Female Mouse BAC"
1. .1091
/note="assembly-fragment"
misc_feature
1192. .3065
/note="assembly-fragment"
misc_feature
3166. .4305
/note="assembly-fragment"
misc_feature
4406. .5983
/note="assembly-fragment"
misc_feature
6084. .8322
/note="assembly-fragment"
misc_feature
8423. .10098
/note="assembly-fragment"
misc_feature
10199. .12592
/note="assembly-fragment"
misc_feature
12693. .14684
/note="assembly-fragment"
misc_feature
14785. .17163
/note="assembly-fragment"
misc_feature
17264. .19900
/note="assembly-fragment"
misc_feature
20001. .23054
/note="assembly-fragment"
misc_feature
23155. .26011
/note="assembly-fragment"
misc_feature
29413. .29512
/note="assembly-fragment"
misc_feature
32991. .3479
/note="assembly-fragment"
misc_feature
33091. .36913
/note="assembly-fragment"
misc_feature
37014. .40880
/note="assembly-fragment"
misc_feature
40981. .45015
/note="assembly-fragment"
misc_feature
45116. .50321
/note="assembly-fragment"
misc_feature
50422. .54284
/note="assembly-fragment"
misc_feature
54385. .59430
/note="assembly-fragment"
misc_feature
59531. .64373
/note="assembly-fragment"
misc_feature
64474. .70503
/note="assembly-fragment"
misc_feature
70604. .77829
/note="assembly-fragment"
misc_feature
77939. .85976
/note="assembly-fragment"
misc_feature
86077. .91694
/note="assembly-fragment"
misc_feature
91795. .99208
/note="assembly-fragment"
misc_feature
99309. .107499
/note="assembly-fragment"
misc_feature


```

/note="assembly_fragment"
misc_feature      107600..115731
/note="assembly_fragment"
misc_feature      115832..123297
/note="assembly_fragment
clone_end:T7

```

Query Match 16.2%; Score 372.2; DB 62; Length 200727;
 Best Local Similarity 85.9%; Pred. No. 7.7e-62;
 Matches 413; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

OY 233 agggcgagcctccaccaccctttagagacaagaagcgtgccgaatagctgcgcgca 292
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150286 AggctGcAGCCCTCAGCAGATCTCTCTCCGAGATGTCGAGTACAGCTGCCGGA 150345

OY 293 gctgacttcaaccgctacgtgacccgaltggccgtgcccgaagcgaagccggtcacca 352
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
150346 GCTGCACTACACCCGCTTCTCTGACAGAGAGGCGCCATGCCGAGCGCCAGCGGTCACGA 150405

      353 gctggtgtgctcgcgccaagtgcgagcccgccgctgctgcccacgcacatcgccgcgca 412
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150406 GTTGTGTGCTCCGCGCAGTCCGCGCCCGCGCGCTCTGCTCCACACGCCATCGGCGCGT 150465

OY 413 caagtgtgagcaccatgtgagccgacttcgctgcattcccgaccgctacgcgcgca 472
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150466 GAAGTGTGTGCGCGCGCGAGACCGGATTTCCGCTGCAATCCGATCGTACCGCGCGCA 150525

OY 473 ggcggtgacgctgctgtgtcccggtggtgagcgccgcgcgcgaaggtgcgctggt 532
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150526 GCGGGTGTGAGCTGTGTGCGCGCGCGCGCGCGCGCTCGCGCAAGTGTGCTGCT 150585

OY 533 ggcctcgtgcaagtgcgaagcgccctcaccgctccacaaccagtgcgaagctcaagaact 592
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150586 GGCTCTGTGCAAGTGTCAAGCGCTTCAACCGCTTCCACACCACTCGAGCTCAAGGACTT 150645

OY 593 cgggaccgagcgctgcgcgcgaagaaggcggaaagccgagcccgccggaagcgc 652
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150646 CGGCGCGGAGACCGCGCGCGCGCGAGAGGGTCCGCAAGCCCGCGCGCGCGCGGAGC 150705

OY 653 caagcccaaccagcgcgagctgaggaagcgtactagagccgcgcgcgcctcccaac 712
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150706 CAAAGCCACACGCGAGCTGTGAGAACGCTTACTAGAGCGAGCCCGCGCTTATGCAGCC 150765

OY 713 c 713
      |
150766 C 150766

```

Search completed: November 9, 2001, 00:00:48
 Job time: 16818 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:31 ; Search time 5549.44 Seconds
(without alignments)
906.203 Million cell updates/sec

Title: US-09-668-021-15

Perfect score: 532

Sequence: 1 agaatgataccacagaatc.....accaagccagccggccga 532

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Archived: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:27 ; Search time 5549.44 Seconds
(Without alignments)
1148.084 Million cell updates/sec

Title: US-09-668-021-13
Perfect score: 674
Sequence: 1 gagagacgcgagtccttccct.....agctgagagaacgcctactag 674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Archived: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est148:*
 118: gb_est149:*
 119: gb_est150:*
 120: gb_est151:*
 121: gb_est152:*
 122: gb_est153:*
 123: gb_est154:*
 124: gb_est155:*
 125: gb_est156:*
 126: gb_est157:*
 127: gb_est158:*
 128: gb_est159:*
 129: gb_est160:*
 130: gb_est161:*
 131: gb_est162:*
 132: gb_est163:*
 133: gb_est164:*
 134: gb_est165:*
 135: gb_est166:*
 136: gb_est167:*
 137: gb_est168:*
 138: gb_est169:*
 139: gb_est170:*
 140: gb_est171:*
 141: gb_est172:*
 142: gb_est173:*
 143: gb_est174:*
 144: gb_est175:*
 145: gb_est176:*
 146: gb_est177:*
 147: gb_est178:*
 148: gb_est179:*
 149: gb_est180:*
 150: gb_est181:*
 151: gb_est182:*
 152: gb_est183:*
 153: gb_est184:*
 154: gb_est185:*
 155: gb_est186:*
 156: gb_est187:*
 157: gb_est188:*
 158: gb_est189:*
 159: gb_est190:*
 160: gb_est191:*
 161: gb_est192:*
 162: gb_est193:*
 163: gb_est194:*
 164: gb_est195:*
 165: gb_est196:*
 166: gb_est197:*
 167: gb_est198:*
 168: gb_est199:*
 169: gb_est200:*
 170: gb_est201:*
 171: gb_est202:*
 172: gb_est203:*
 173: gb_est204:*
 174: gb_est205:*
 175: gb_est206:*
 176: gb_est207:*
 177: gb_est208:*
 178: gb_est209:*
 179: gb_est210:*
 180: gb_est211:*
 181: gb_est212:*
 182: gb_est213:*
 183: gb_est214:*
 184: gb_est215:*
 185: gb_est216:*
 186: gb_est217:*
 187: gb_est218:*
 188: gb_est219:*
 189: gb_est220:*

190: gb_est110:*
 191: gb_est111:*
 192: gb_est112:*
 193: gb_est113:*
 194: gb_est114:*
 195: gb_est115:*
 196: gb_est116:*
 197: gb_est117:*
 198: gb_est118:*
 199: gb_est119:*
 200: gb_est120:*
 201: gb_est121:*
 202: gb_est122:*
 203: gb_est123:*
 204: gb_est124:*
 205: gb_est125:*
 206: gb_est126:*
 207: gb_est127:*
 208: gb_est128:*
 209: gb_est129:*
 210: gb_est130:*
 211: gb_est131:*
 212: gb_est132:*
 213: gb_est133:*
 214: gb_est134:*
 215: gb_est135:*
 216: gb_est136:*
 217: gb_est137:*
 218: gb_est138:*
 219: gb_est139:*
 220: gb_est140:*
 221: gb_est141:*
 222: gb_est142:*
 223: gb_est143:*
 224: gb_est144:*
 225: gb_est145:*
 226: gb_est146:*
 227: gb_est147:*
 228: gb_est148:*
 229: gb_est149:*
 230: gb_est150:*
 231: gb_est151:*
 232: gb_est152:*
 233: gb_est153:*
 234: gb_est154:*
 235: gb_est155:*
 236: gb_est156:*
 237: gb_est157:*
 238: gb_est158:*
 239: gb_est159:*
 240: gb_est160:*
 241: gb_est161:*
 242: gb_est162:*
 243: gb_est163:*
 244: gb_est164:*
 245: gb_est165:*
 246: gb_est166:*
 247: gb_est167:*
 248: gb_est168:*
 249: gb_est169:*
 250: gb_est170:*
 251: gb_est171:*
 252: gb_est172:*
 253: gb_est173:*
 254: gb_est174:*
 255: gb_est175:*
 256: gb_est176:*
 257: gb_est177:*
 258: gb_est178:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:24 ; Search time 5549.44 Seconds
(without alignments)
1086.762 Million cell updates/sec

Title: US-09-668-021-11
Perfect score: 638

Sequence: 1 atgcagccctcactagaccc.....ctgagaaagcctactagag 638

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Archived: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:19 ; Search time 5549.44 Seconds
(without alignments)
1093.576 Million cell updates/sec

Title: US-09-668-021-9
Perfect score: 642
Sequence: 1 atgcagctccactggccct.....agctggagagcgcctactag 642

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Database: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*

190: gb_est110:*
191: gb_est111:*
192: gb_hic:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:10:40 ; Search time 9526.57 Seconds
(without alignments)
863.778 Million cell updates/sec

Title: US-09-668-021-15
Perfect score: 532
Sequence: 1 agaatgatgcacagaatc.....accgaagccagcgagccga 532

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Archived: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_ov:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pi1:*
- 13: gb_pi2:*
- 14: gb_pi3:*
- 15: gb_pi4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_hrgo_hum:*
- 20: em_hrgo_inv:*
- 21: em_hrgo_rod:*
- 22: em_hrg_hum1:*
- 23: em_hrg_hum2:*
- 24: em_hrg_hum3:*
- 25: em_hrg_hum4:*
- 26: em_hrg_hum5:*
- 27: em_hrg_hum6:*
- 28: em_hrg_hum7:*
- 29: em_hrg_hum8:*
- 30: em_hrg_inv1:*
- 31: em_hrg_inv2:*
- 32: em_hrg_other:*
- 33: em_hrg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_ov:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pi:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_v11:*
- 59: gb_v12:*
- 60: gb_v13:*
- 61: gb_v14:*
- 62: gb_v15:*
- 63: gb_v16:*
- 64: gb_v17:*
- 65: gb_v18:*
- 66: gb_v19:*
- 67: gb_v20:*
- 68: gb_v21:*
- 69: gb_v22:*
- 70: gb_v23:*
- 71: gb_v24:*
- 72: gb_v25:*
- 73: gb_v26:*
- 74: gb_v27:*
- 75: gb_v28:*
- 76: gb_v29:*
- 77: gb_v30:*
- 78: gb_v31:*
- 79: gb_v32:*
- 80: gb_v33:*
- 81: gb_v34:*
- 82: gb_v35:*
- 83: gb_v36:*
- 84: gb_v37:*
- 85: gb_v38:*
- 86: gb_v39:*
- 87: gb_v40:*
- 88: gb_v41:*
- 89: gb_v42:*
- 90: gb_v43:*
- 91: gb_v44:*
- 92: gb_v45:*
- 93: gb_v46:*
- 94: gb_v47:*
- 95: gb_v48:*
- 96: gb_v49:*
- 97: gb_v50:*
- 98: em_ba3:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	532	100.0	7 AF326738	AF326738 Bos tauru
2	467.8	87.9	2236 AF331844	AF331844 Homo sapi
3	467.8	87.9	2323 AF326739	AF326739 Homo sapi
4	467.8	87.9	2329 AX056687	AX056687 Sequence
5	461.4	86.7	642 89 AF326742	AF326742 Cercopit
6	424.6	79.8	674 94 AF326741	AF326741 Rattus no
7	414.2	77.9	638 94 AF326740	AF326740 Mus muscu
8	375.2	70.5	21501 89 AF326736	AF326736 Homo sapi

C	9	375.2	70.5	94752	85	AC003098	AC003098 Homo sapi
C	10	375.2	70.5	151780	72	AC055813	AC055813 Homo sapi
C	11	332.4	62.5	117423	94	AF336737	AF336737 Mus muscu
C	12	332.4	62.5	110003	73	AC068782	AC068782 Mus muscu
C	13	332.4	62.5	200727	62	AC012226	AC012226 Mus muscu
C	14	332.4	62.5	211533	73	AC068807	AC068807 Mus muscu
C	15	288.4	54.2	177143	75	AC073954	AC073954 Homo sapi
C	16	86.4	16.2	51575	68	AC023810	AC023810 Mus muscu
C	17	72.2	13.6	900	9	AR017832	AR017832 Sequence
C	18	66	12.4	150951	93	AC079155	AC079155 Homo sapi
C	19	64.4	12.1	1500	93	HSMB00108	AL050024 Homo sapi
C	20	61	11.5	412	10	AX071502	AX071502 Sequence
C	21	59	11.5	204224	12	AC087881	AC087881 Mus muscu
C	22	61	11.1	1153	53	CNS07BPV	AL438395 T3 end of
C	23	58.4	11.0	165448	67	AC010827	AC010827 Homo sapi
C	24	56.8	10.7	63082	61	AC022663	AC022663 Homo sapi
C	25	55.8	10.5	252859	68	AC024907	AC024907 Homo sapi
C	26	55.6	10.5	37856	2	AX024319	AX024319 Sequence
C	27	55.6	10.5	37856	9	AX024212	AX024212 Sequence
C	28	55.4	10.4	1153	53	CNS07BPV	AL438395 T3 end of
C	29	54	10.2	187738	71	AC027682	AC027682 Homo sapi
C	30	53.2	10.0	52884	73	AC068018	AC068018 Homo sapi
C	31	53	10.0	38366	3	SC23B6	AL390686 Streptomy
C	32	52.8	9.9	70097	68	AC033550	AC033550 Homo sapi
C	33	52.2	9.8	162005	80	AL356002	AL356002 Homo sapi
C	34	52	9.8	567	53	CNS075A1	AL429856 Clone BA0
C	35	52	9.7	2223	3	KNA133131	ALJ33331 Myxococcu
C	36	51.8	9.7	35624	3	SCB69	AL079308 Streptomy
C	37	51.8	9.7	51053	64	AC016154	AC016154 Homo sapi
C	38	51.8	9.7	174424	97	HSU52112	US2112 Homo sapien
C	39	51.4	9.7	10948	1	AE005746	AE005746 Caldicellu
C	40	51.4	9.6	720	3	TTN409330	AL409330 T. thermus t
C	41	51	9.6	181191	75	AC080138	AC080138 Mus muscu
C	42	51	9.6	278652	75	AC073809	AC073809 Mus muscu
C	43	50.6	9.5	63681	78	AC090582	AC090582 Homo sapi
C	44	50.4	9.5	12778	1	AE005877	AE005877 Caldicellu
C	45	50.2	9.4	52884	73	AC068018	AC068018 Homo sapi

ALIGNMENTS

	RESULT	1			
	AF326738				
LOCUS	AF326738	532 bp	mRNA	MM	28-FEB-2001
DEFINITION	Bos taurus sclerostin mRNA, partial cds.				
ACCESSION	AF326738				
VERSION	AF326738.1	GI:13161016			
KEYWORDS	.				
SOURCE	COW.				
ORGANISM	Bos taurus				
	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
	Bovidae; Bovinae; Bos.				
REFERENCE	1 (bases 1 to 532)				
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepier,B.W., Koyaveich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Aitsch,R.S., Gillett,L.,Colbert,T., Tacconi,P., Galas,D. Hamersma,H., Beignton,P. and Mulligan,D.T. Bone Dysplasia Cytosteosis Results from Loss of the SOST Gene Product, a Novel Cyclase Knot-Containing Protein Am. J. Hum. Genet. 68 (3), 577-589 (2001) 11179006				
TITLE	2 (bases 1 to 532)				
JOURNAL	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepier,B.W., Koyaveich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Aitsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D. Hamersma,H., Beignton,P. and Mulligan,D.T. Direct Submission Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA				
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
FEATURES					
source	Location/Qualifiers				. .532

```

CDS
        /organism="Bos taurus"
        /db_xref="taxon:9913"
<1...>532
        /codon_start=3
        /product="sclerostin"
        /protein_18="AAK13453.1"
        /db_xref="GI:13161017"
        /translation="NDATFTPELGEYPEPLPELNNKTYMNRANGSRPHHPETTKDA
        SYSSCSLEHFRVYTDPCRSKAPVTELVSQCGARLLPAIGRGWMPSPGDFR
        CIPDYSRRQRVOLCLPGCAAPRAKRVLAIVSCSKRLTRFHNOSELKDFGEARPQT
        GKLRPRARGTAKSRA"
BASE COUNT      95 a      199 c      176 g      62 t
ORIGIN
Query Match      100.0%; Score 532; DB 7; Length 532;
Best Local Similarity 100.0%; Pred. No. 5e-77;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 agaatgtagtccaaagaatcatcatcccgagctcgggcgagatcaccccgagccctctgcagagc 60
    |||||||
Db 1 AGAATTGTCGACAGAAATCATCTCCCGAGCTGGGCGAGTACCCGAGACCTCTCCAGAGC 60

OY 61 tgaacaacaagaccatgaaacccggcgagaaacggagaggagacctcccaaccaccttgg 120
    |||||||
Db 61 TGAACAACAAGACCAATGAACCCGGGCGAGAAAGCAAGAGACACCTCCACACCCCTTTGG 120

OY 121 agaccaaagagcgtctcgagtagtacaactctcgggagctgcaacttcaaccgcgtacgtgacgg 180
    |||||||
Db 121 AGACCAAGAAGCGCTCTCGAGTACAGCTGCGGAGACTGTCACTTACCCCGTACGTACCG 180

OY 181 atgggcctgtcgagcgccaaagccggttcaaccagctgtgtgtctcgggccaagtctgcgc 240
    |||||||
Db 181 ATGGGCGGTGTCGCCACAGCGCAAGCGGTACCGAGCTGTGTGTCTCGGCGCAATGCGGCC 240

OY 241 cggcgccctctgtgcgccaacgcacacggcgccgcaagtgtgtgtgtgcgccaagcgggccgg 300
    |||||||
Db 241 CGGCGCGCTCTGTGCCAAGCGCATGGGCGCGGAATGTGTGGCGGCCCAAGCGGGCCGG 300

OY 301 acttcgcgtgcatcccccagacgctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
    |||||||
Db 301 ACTTCCGCTCTCATCCCGACCGCTTACCGCGCGACCGGGTCCACCTGTGTGTCTGTGCGCG 360

OY 361 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
    |||||||
Db 361 GCGGCGCGCGCGCGCGCGCGCAAGTGTGCGCTGTGTGCGCTGTGCAATGTCAAGCGCTCA 420

OY 421 ctgcgtctcaacaaccagtcgcagctcaagaagacttcgggcccgcgcgcgcgcgcgcgcgcgc 480
    |||||||
Db 421 CTCGCTTCCCAACACAGTCTCGAGCTCAAGGACTTGTGGGCCCGGAGCGCGCGCGCCCAAA 480

OY 481 cggcgccggaagctctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 532
    |||||||
Db 481 CGGCGCGGAAGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 532

RESULT 2
AF331844 2296 bp mRNA PRI 06-MAR-2001
LOCUS AF331844 Homo sapiens SOST (SOST) mRNA, complete cds.
DEFINITION AF331844
ACCESSION AF331844
VERSION AF331844.1 GI:13236417
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 2296)
Bailemans,W., Ebeling,M., Patel,N., Van Hul,E., Olson,P.,
Dlostedt,M., Iacra,C., Wuyts,W., van den Ende,J., Willems,P.,
Praes-Ayres,A.F., Hill,S., Bueno,M., Ramos,F.J., Taccori,B.,
Dikkers,F.G., Stratakis,C., Lindpaintner,K., Vickyry,B.,
Foerzler,F., and Van Hul,W.

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:07:17 ; Search time 9526.57 Seconds

(without alignments)
1094.335 Million cell updates/sec

Title: US-09-668-021-13

Perfect score: 674

Sequence: 1 gagagccgagtcctctct.....agctgagagcgcctactag 674

Scoring table: IDENTITY_NDC

Gapop 10.0 , Gapext 1.0

arched: 1344157 segs, 7733874588 residues 2688314

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pi1:*
13: gb_pi2:*
14: gb_pi3:*
15: gb_pi4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_hcgo_hum:*
20: em_hcgo_inv:*
21: em_hcgo_rtd:*
22: em_hcgo_hum1:*
23: em_hcgo_hum2:*
24: em_hcgo_hum3:*
25: em_hcgo_hum4:*
26: em_hcgo_hum5:*
27: em_hcgo_hum6:*
28: em_hcgo_hum7:*
29: em_hcgo_hum8:*
30: em_hcgo_inv1:*
31: em_hcgo_inv2:*
32: em_hcgo_other:*
33: em_hcgo_rtd:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pi:*
48: em_pi1:*
49: em_pi2:*
50: em_pi3:*
51: em_pi4:*
52: em_pi5:*
53: em_pi6:*
54: em_pi7:*
55: em_pi8:*
56: em_pi9:*
57: em_pi10:*
58: em_pi11:*
59: em_pi12:*
60: em_pi13:*
61: em_pi14:*
62: em_pi15:*
63: em_pi16:*
64: em_pi17:*
65: em_pi18:*
66: em_pi19:*
67: em_pi20:*
68: em_pi21:*
69: em_pi22:*
70: em_pi23:*
71: em_pi24:*
72: em_pi25:*
73: em_pi26:*
74: em_pi27:*
75: em_pi28:*
76: em_pi29:*
77: em_pi30:*
78: em_pi31:*
79: em_pi32:*
80: em_pi33:*
81: em_pi34:*
82: em_pi35:*
83: em_pi36:*
84: em_pi37:*
85: em_pi38:*
86: em_pi39:*
87: em_pi40:*
88: em_pi41:*
89: em_pi42:*
90: em_pi43:*
91: em_pi44:*
92: em_pi45:*
93: em_pi46:*
94: em_pi47:*
95: em_pi48:*
96: em_pi49:*
97: em_pi50:*
98: em_pi51:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	674	100.0	674	94	AF326741 Rattus no
2	556	82.5	638	94	AF326740 Mus muscu
3	536.2	79.6	2296	89	AF331844 Homo sapi
4	536.2	79.6	2293	89	AF326739 Homo sapi
5	536.2	79.6	2329	9	AX056687 Sequence
6	530	78.6	642	89	AF326742 Cercopit
7	424.6	63.0	532	7	AF326738 Bos tauru
8	407.4	60.4	17423	94	AF326737 Mus muscu

ALIGNMENTS

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:04:04 : Search time 9526.57 Seconds
(without alignments)
1035.884 Million cell updates/sec

Title: US-09-668-021-11

Perfect score: 638
Sequence: 1 atgagccctactagcccc.....ctggagaacgctactagag 638

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

arched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
 - 1: gb_ba1:*
 - 2: gb_ba2:*
 - 3: gb_ba3:*
 - 4: gb_in1:*
 - 5: gb_in2:*
 - 6: gb_in3:*
 - 7: gb_om:*
 - 8: gb_ov:*
 - 9: gb_pat1:*
 - 10: gb_pat2:*
 - 11: gb_ph:*
 - 12: gb_pl1:*
 - 13: gb_pl2:*
 - 14: gb_pl3:*
 - 15: gb_pl4:*
 - 16: em_ba1:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_higo_hum:*
 - 20: em_higo_inv:*
 - 21: em_higo_rod:*
 - 22: em_hig_hum1:*
 - 23: em_hig_hum2:*
 - 24: em_hig_hum3:*
 - 25: em_hig_hum4:*
 - 26: em_hig_hum5:*
 - 27: em_hig_hum6:*
 - 28: em_hig_hum7:*
 - 29: em_hig_hum8:*
 - 30: em_hig_hum11:*
 - 31: em_hig_inv2:*
 - 32: em_hig_other:*
 - 33: em_hig_rod:*
 - 34: em_hum1:*
 - 35: em_hum2:*
 - 36: em_hum3:*
 - 37: em_hum4:*
 - 38: em_hum5:*
 - 39: em_hum6:*
 - 40: em_hum7:*
 - 41: em_in:*
 - 42: em_om:*
 - 43: em_or:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	638	100.0	638	94 AF326740	AF326740 Mus muscu
2	556	87.1	674	94 AF326741	AF326741 Rattus no
3	503.2	78.9	642	89 AF326742	AF326742 Cercopith
4	502	78.7	2296	89 AF331844	AF331844 Homo sapi
5	502	78.7	2323	89 AF326739	AF326739 Homo sapi
6	502	78.7	2329	9 AX056687	AX056687 Sequence
7	427	66.9	17423	94 AF326737	AF326737 Mus muscu
8	427	66.9	110000	73 AC068782_2	Continuation (3 of

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAGAGTGTGTCTCGATACACTCCCGCAG	240
Query	241	ctgagactaacccgcctctctgacagagcgccatgpcgcgagcgccaaagccggtacacgag	300

BASE COUNT	114	a	224	c	209	g	91	t
ORIGIN								

CDS	1..636	/codon_start=1	/product="sclerostin"	/protein_id="AAK13455.1"	/db_xref="taxon:10090"	/chromosome="11"

Query Match	100.0%;	Score 638;	DB 94;	Length 638;
Best Local Similarity	100.0%;	Prod. NO. 6.1e-101;		
Matches 638; Conservative	0;	Mismatches	0;	Indels
			0;	Gaps
				0;

Query	1	atgagcctcaatagccccggtgcatactgcactctgtgcagctgctctgtgct	60
Db	1	ATGCAGCCCTCATACACCCCGCTGCCTCATTCGCTACTGTCGACGCTGCTTGTCT	60
Query	61	gtgagagggccagggtgggaagccttgagatgatccaaagttatcccaaggctt	120
Db	61	GTGGAGGGCCAGGGGTGGCAAGCTTCAGAGATGTGTCACAGAGGTCATCCAGGGCTT	120
Query	121	ggagagtaaccgagagctctctctgtagaaacaacagacatgaaccgggaggaatgga	180
Db	121	GGAAGTACCCCGAGGCTCTCTCTGAGAAACAACAGACATGAACCGGCGCAGATGGA	180
Query	181	ggcagacatcccaatctccatagacccaagaagtgtgtccagatagagcttgcgcgaag	240
Db	181	GGCAGACCTCCACCATCTCTATAGCCCAAG	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:00:48 ; Search time 9526.57 Seconds
(without alignments)
1042.379 Million cell updates/sec

Title: US-09-668-021-9

Perfect score: 642

Sequence: 1 atgcagctccactgcctct.....agctgagagacgcctactag 642

Scoring table: IDENTITY_NTC

Gapop 10.0, Gapext 1.0

arched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

GenBml:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_om:*

9: gb_ov:*

10: gb_pat1:*

11: gb_pat2:*

12: gb_ph:*

13: gb_pi1:*

14: gb_pi2:*

15: gb_pi3:*

16: gb_pi4:*

17: em_ba1:*

18: em_ba2:*

19: em_fun:*

20: em_htgo_hum:*

21: em_htgo_inv:*

22: em_htgo_rod:*

23: em_htg_hum1:*

24: em_htg_hum2:*

25: em_htg_hum3:*

26: em_htg_hum4:*

27: em_htg_hum5:*

28: em_htg_hum6:*

29: em_htg_hum7:*

30: em_htg_hum8:*

31: em_htg_inv1:*

32: em_htg_inv2:*

33: em_htg_rod:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_om:*

43: em_or:*

Result No.	Score	Query Match	Length	DB	ID	Description
1	642	100.0	642	89	AF326742	Cercopit
2	613.2	95.5	2296	89	AF331844	Homo sapi
3	613.2	95.5	2323	89	AF326739	Homo sapi
4	613.2	95.5	2329	9	AX056687	Sequence
5	530	82.6	674	94	AF326741	Rattus no
6	503.2	78.4	638	94	AF326740	Mus muscu
7	461.4	71.9	532	7	AF326738	Bos tauru
8	413.4	64.4	21501	89	AF326736	Homo sapi

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	9	413.4	94752	85	AC003098	AC003098 Homo sapi	
C	10	413.4	151780	72	AC055813	AC055813 Homo sapi	
C	11	361	56.2	17423	94	AF326737	AF326737 Mus muscu
C	12	361	56.2	110000	73	AC068782	Continuation (3 o
C	13	361	56.2	200727	62	AC012296	AC012296 Mus muscu
C	14	361	56.2	211533	73	AC068807	AC068807 Mus muscu
C	15	324.6	50.6	177744	75	AC073954	AC073954 Homo sapi
C	16	146.8	22.9	51575	68	AC023810	AC023810 Mus muscu
C	17	71	900	9	AR017832	AR017832 Sequence	
C	18	64.4	10.0	150951	75	AC079155	AC079155 Homo sapi
C	19	62.8	9.8	1500	93	HSM890108	AL050024 Homo sapi
C	20	57.8	9.0	412	10	AX071502	AX071502 Sequence
C	21	57.8	9.0	204224	77	AC087891	AC087891 Mus muscu
C	22	54.2	8.4	201586	83	CNS018D91	AL136000 Homo sapi
C	23	54	8.4	37856	2	AX024319	AX024319 Sequence
C	24	54	8.4	37856	9	AX024212	AX024212 Sequence
C	25	52.4	8.2	207091	77	AC087289	AC087289 Homo sapi
C	26	52.2	8.1	2943	10	117281	117281 Sequence
C	27	52.2	8.1	2943	59	HSV8Q	M57388 Simian hea
C	28	52.2	8.1	4585	59	HVS8GB	X56935 Herpesvitr
C	29	51.6	8.0	1153	53	CNS07BVP	AL438395 T3 end O
C	30	51.4	8.0	229380	76	AC079636	AC079636 Mus muscu
C	31	51.2	8.0	70097	68	AC023530	AC023530 Homo sapi
C	32	50.6	7.9	52884	73	AC068018	AC068018 Homo sapi
C	33	50.4	7.9	165448	61	AC010827	AC010827 Homo sapi
C	34	50.4	7.9	187738	71	AC027682	AC027682 Homo sapi
C	35	50.2	7.8	1133	53	CNS07BVP	AL438395 T3 end O
C	36	49.6	7.7	175837	67	AC022045	AC022045 Homo sapi
C	37	49.6	7.7	215225	91	AP001972	AP001972 Homo sapi
C	38	49.4	7.7	51053	64	AC016154	AC016154 Homo sapi
C	39	49.4	7.7	181405	87	AC018682	AC018682 Homo sapi
C	40	49.4	7.7	252859	68	AC024907	AC024907 Homo sapi
C	41	49.2	7.7	31636	3	SC7C7	AL031031 Streptomy
C	42	48.9	7.6	96745	75	AC074219	AC074219 Mus muscu
C	43	48.8	7.6	147094	62	AC012027	AC012027 Homo sapi
C	44	48.6	7.6	9912	1	AE001883	AE001883 Deinococc
C	45	48.6	40745	3	SC151	AL109548 Streptomy	

ALIGNMENTS

AC003098	Homo sapi
AC005583	Homo sapi
AE326737	Mus muscu
Continuation (3 of 3)	
AC012266	Mus muscu
AC068807	Mus muscu
AC073954	Homo sapi
AC023810	Mus muscu
AR017832 Sequence	
AC079155	Homo sapi
AL050024	Homo sapi
AX071502	Sequence
AC087891	Mus muscu
AL136600	Homo sapi
AX024313	Sequence
AX024212	Sequence
AC087289	Homo sapi
I17281 Sequence 7	
M57388	Simian hefte
X55635	Therapsid musc
AL383335	T3 end of
AC079636	Mus muscu
AC023530	Homo sapi
AC068018	Homo sapi
AC010827	Homo sapi
AC027652	Homo sapi
AL383335	T3 end of
AC022045	Homo sapi
AP001192	Homo sapi
AC016154	Homo sapi
AC018682	Homo sapi
AC024907	Homo sapi
AL031031 Streptomy	
AC074219	Mus muscu
AC012027	Homo sapi
AE001883	Dicnocc
AL109848	Streptomy

	RESULT	1			
	AF326742				
LOCUS	AF326742	642 bp	mRNA	PRI	28-FEB-2001
DEFINITION	Cercopithecus aethiops sclerostin mRNA, complete cds.				
ACCESSION	AF326742				
VERSION	AF326742.1	GI:13161028			
KEYWORDS	.	African green monkey.			
ORGANISM	JURCE	Cercopithecus aethiops.			
		Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;			
		Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
		Cercopithecinae; Cercopithecus.			
REFERENCE	1 (bases 1 to 642)				
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepfer,B.W., Kovachewich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Ailtsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P., and Mulligan,J.T. Bone Dysplasia Scleroosteosis Results from Loss of the SOSR Gene Product, a Novel Cystine Knot-Containing Protein <i>Am. J. Hum. Genet.</i> 68 (3), 577-589 (2001) 11179006				
TITLE	2 (bases 1 to 642)				
JOURNAL	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepfer,B.W., Kovachewich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Ailtsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P., and Mulligan,J.T. Direct Submission Submitted (07-DEC-2000) Genomics, Celleghe Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA				
PUBMED					
REFERENCE					
AUTHORS					
FEATURES	source	I..642			
	location/Qualifiers				

CDS	/organism="Cercopithecus aethiops" /db_xref="taxon:9534" 1..642 /codon_start=1 /product="sclerostin" /protein_id="AAK13457.1" /db_xref="gi:13161029" /translation="MQLPLALCTLVLAHAFRVEGSGMQAFKKDATEIIPELGEYR EPPPLENNKTMMNRANGRRPHHPPTKPSYSRRLHPLRYTLDGCSAKRVTFE LYSQCCPARLLPNAIGCKRMRRPSPSPDRCTIPDXRYROROLLCPGGAAPRAKRV LVASKCKRLTRFHNOSELKDFEAPARPPKRRPPRRARGAAMAELENAI"									
BASE COUNT	113	a	228	c	210	g	91	t		
ORIGIN										
Query Match	100.0%	Score 642:	DB 89:	Length 642:						
Best Local Similarity	100.0%	Pred. No. 4.1e-102:								
Matches 642:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:		
Db	1	ATCGAGCTCCACATGGAGCCCTGTGTCTTGTGTCCTCGCTGATACACGACGCTTCCGHTA	60							
Oy	1	atcgagctccacatgagccctgtgtcttgtcttgcctgtctgtatcacagagcttcctgta	60							
Db	1	ATCGAGCTCCACATGGAGCCCTGTGTCTTGTGTCCTCGCTGATACACGACGCTTCCGHTA	60							
Oy	61	gtgagagcgccagaggtgtgcaagcctcaagaatgtgtcccaagaaatcatcccgaggtc	120							
Db	61	GTGAGAGCGCCAGGAGGCGGACAGGCGCTTCAAGATGTGTGCACGGAATATCTCCCGAGTTC	120							
Oy	121	ggaagatgaccccgagctccacccgagctctgagaaacaagaacatgaaacgagcgagag	180							
Db	121	GGAGAGTACCCCAAGCCTCCACCGAGCTCGAGACGAGAACAAAGACCATGAAACGGGCGGAG	180							
Oy	181	aatgagagcgagcctcccaaccaacccctttagaccaaagaagtgctcagatacagctgc	240							
Db	181	AATGAGAGGCGGCGCTCCACACACCCCTTTGAGACAAAGACGGTCCGATACAGTCG	240							
Oy	241	cgaagctgtcaacttaaccgcgaactgtgacaggaatggccgttcgcaagcgcaagccagtc	300							
Db	241	CGAAGCTGCACATTCACCCGCTACTGTGACCCGATGGGCGGTGCCGACGCCAAGCCAGTC	300							
Oy	301	accgaagctgtgtgtctcccgagctgagcgccgagcaagctgtctgtgccaaagccatctgc	360							
Db	301	ACCGAGTTGGTGTGTCTCCGCGCCAGTCTGGGCCCGGACACCCCTGCTGCCAACGCCATCTGC	360							
Oy	361	cgcgagcaagtgatgtggcccgagatggagccgaattccgcgtgcatcccgagccgactacgc	420							
Db	361	CGCGCAAGTGTGTGGCCCGGAGTGGGCCCCGACTTCCCTGTGATCCCGACCGCTACGCG	420							
Oy	421	gcgcagagctltgcaagctgt	480							
Db	421	GCGCAGAGGTGTGACGCTGTGTCTCCGGTGGTGTGCGCGCGCGCGCGCAAGTGTGCG	480							
Oy	481	ctgtgtgctctgtgcaagtgcaagtgcaagcgctcaaccgctcttcacaacacagtgtagactaa	540							
Db	481	CTGTGTGCTCTGTGTCAAGTGTCAAGCGCTTCAACCGCTTCCACAAACAGTGTGAGACTTAAG	540							
Oy	541	gacttggtcccgagagcgcgcttgagccgagaaagcgagaaacgagcccgcccgag	600							
Db	541	GACTTGCGTCCGAGGCGCGCTCGCGCGGAGAAAGGGCGCGGAGCGCGCGCGCGCGG	600							
Oy	601	ggggccaaagccaatcagagccagacttgagaagcgctactag	642							
Db	601	GGGGCCAAAGCCAATCAGGCGCGAGCTGTGGAGAACGCGCTACTAG	642							
RESULT	2									
AF331844	AF331844	2296 bp	mrna	PRI						
LOCUS	AF331844									
DEFINITION	Homo sapiens SOST (SOST) mRNA, complete cds.									
ACCESSION	AF331844									
VERSION	AF331844.1	GI:13236417								
KEYWORDS	human.									
SOURCE	human.									
ORGANISM	Homo sapiens									